



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**DETECCIÓN DE ESTILOS DE JUEGO EN FÚTBOL PROFESIONAL
MEDIANTE INFERENCIA BAYESIANA**

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

CONSTANZA PATRICIA ENCINA VILLEGAS

PROFESOR GUÍA:
DENIS SAURÉ VALENZUELA

MIEMBROS DE LA COMISIÓN:
ANDRÉS MUSALEM SAID
GUILLERMO ALFREDO DURÁN
MARIO GUAJARDO ANDRADES

Este trabajo ha sido parcialmente financiado por:
Proyecto Anid Fondecyt 1211407

SANTIAGO DE CHILE
2023

RESUMEN DE TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN GESTIÓN DE OPERACIONES Y
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA
CIVIL INDUSTRIAL
POR: CONSTANZA PATRICIA ENCINA VILLEGAS
FECHA: 2023
PROFESOR GUÍA: DENIS SAURÉ VALENZUELA

DETECCIÓN DE ESTILOS DE JUEGO EN FÚTBOL PROFESIONAL MEDIANTE INFERENCIA BAYESIANA

El análisis de fútbol cubre una amplia gama de técnicas y aplicaciones, que incluyen seguimiento de jugadores, análisis de rendimiento, optimización de estrategias de juego y reclutamiento. En cuanto al reclutamiento, los clubes profesionales buscan incorporar nuevos jugadores que puedan contribuir a su estrategia de juego para la nueva temporada. Actualmente, traen jugadores que triunfaron en otros clubes, sin tener necesariamente en cuenta la diferencia entre los estilos de juego de los respectivos equipos. Esto plantea la difícil tarea de integrarlos en la dinámica del equipo y adaptarlos al plan de juego.

Es importante para los equipos que estos nuevos fichajes logren adaptarse lo más rápido posible, para que rindan al mismo nivel que en sus clubes anteriores, o mejor aún, que obtengan rendimientos superiores. Al incorporar datos objetivos en el proceso de reclutamiento, los equipos pueden tomar decisiones más informadas y precisas sobre la selección de jugadores, lo que puede tener un impacto significativo en su éxito deportivo y financiero. Por ello, el objetivo de esta investigación es detectar estilos de juego en equipo para identificar cuáles son similares y mejorar la adaptación del nuevo jugador.

Para lograr esto, se propone un método de clasificación no supervisado. En particular, un modelo de inferencia bayesiano, para proporcionar un marco probabilístico para las relaciones entre las métricas de desempeño de cada equipo y los estilos de juego. Incorpora conocimientos previos sobre los datos y ajusta el modelo en consecuencia, lo que genera resultados más precisos y sólidos.

El proyecto utiliza datos InStat del campeonato 2021 de la primera división chilena. Las variables consideradas describen de forma agregada lo ocurrido en el campo a nivel de partido y equipo, tales como: disputas defensivas, disputas ofensivas, pases, tiros al arco, etc.

El modelo bayesiano propuesto estima la distribución de probabilidad de los parámetros “estilo de juego”. Esto permitió, a través de un análisis de los resultados, declarar ciertos “estilos de juego” y qué equipos pertenecen a cada uno con cierta probabilidad.

Este modelo podría extrapolarse potencialmente a muchas ligas, de muchos países y en diferentes deportes de equipo. Esto, para obtener resultados más precisos, facilitar la adaptación de la eficiencia de un jugador a un nuevo equipo o utilizarlo como método de análisis del estilo de juego del equipo rival.

*Para todas las que sueñan,
con ser ingenieras un día.*

Saludos

Agradecimientos

Escribiendo estos agradecimientos, me doy cuenta que el proceso que termina hoy, no partió hace sólo unos meses, sino que ha sido un proceso de poco más de seis años. Cada paso en este camino ha estado marcado por el apoyo incondicional de personas excepcionales, cuya sabiduría ha sido faro en los momentos desafiantes.

Primero que todo, quiero agradecerle a mis papás Patricia y Alejandro, por amarme incondicionalmente, siempre impulsarme a ser mejor y por darme todo lo que necesité y más. A mis hermanos de sangre y alma: Nicolás, por ser mi mejor amigo hoy y siempre, Alejandro, por enseñarme que hay que tomar riesgos y vivir la vida, y Carlos, por quererme, cuidarme y tratarme siempre como tu hermana.

Quiero agradecerle a mis abuelitos Sergio y Angélica, que aunque ya no puedan acompañarme, me enseñaron lo importante que es la familia, y que da lo mismo lo que quiera hacer, lo puedo lograr. A la lolo y el tío Carlos por siempre estar para mí, y por todos los consejos que me dieron y me darán. A todos mis tíos, tías, primos, primas, sobrinas, sobrinos, cuñadas y demás, porque la familia siempre será lo más importante.

A mis familias tutoras 2020, 2021 y 2022 por ser compañía tremenda en tiempos pandémicos y siempre motivarme a seguir aprendiendo. A todos mis amigos universitarios, especialmente al Manuel, la Dani y la Javi, porque si bien los momentos estresantes no faltaron, el apañe tampoco. A la Cami, por recordarme que a tus sueños sólo tú le pones límites. Al Pepe y Mati por darme un abrazo cada vez que lo necesité e incontables historias juntos.

A los CIDS. La Jo y la Nati por empoderarme en días de tesis. Al Benja por infinitas conversaciones en la 601 (la mayoría de las veces con sentido). Al Pablo por ser tremendo compañero y una inspiración diaria.

Quiero agradecerle a todos los equipos docentes de los que fui parte, a la oficina de Docencia de industrias y a todos los profesores que me cruce en este camino, especialmente Andrés Musalem y Denis Sauré.

Y por último, pero no menos importante, quiero agradecerme a mí. Porque la Coni del 8 de marzo de 2017, el día de su inducción, nunca pensó que llegaría donde está, que conocería lugares a los que fue, ni que viviría las experiencias que vivió. Lo hizo a punta de esfuerzo y mucha dedicación.

Me siento inmensamente afortunada y agradecida por haber vivido una experiencia universitaria tremenda. Cada momento ha sido valioso, lleno de aprendizajes y crecimiento personal.

Quiero aprovechar este espacio para animar a cada lector a seguir sus metas. La clave para alcanzar el éxito y la felicidad radica en seguir día a día aquello que te apasiona, sin temor a los desafíos que puedan surgir en el camino. Solo ustedes son los arquitectos de su destino y tienen el poder de decidir hasta dónde llegar.

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivo y Metodología	3
1.3. Contribución	3
2. Revisión de Literatura	4
2.1. Clasificación de estilos de juego	4
2.2. Inferencia Bayesiana	6
3. Datos	8
3.1. Proveedor	8
3.2. EDA	8
3.3. Selección de variables	13
4. Modelo	16
4.1. Modelo probabilista	16
5. Resultados	18
5.1. Monte Carlo Markov Chain	18
5.2. Clasificación estilos de juego	18
5.2.1. Tres estilos de juego	19
5.2.1.1. Estilo 1: Estilo de juego defensivo	19
5.2.1.2. Estilo 2: Estilo de juego ofensivo	20
5.2.1.3. Estilo 3: Estilo de juego aéreo y de construcción	21
5.2.2. Cuatro estilos de juego	22
5.2.2.1. Estilo 1: Estilo defensivo	23
5.2.2.2. Estilo 2: Estilo directo y aéreo	23
5.2.2.3. Estilo 3: Estilo ofensivo de construcción	24
5.2.2.4. Estilo 4: Estilo ofensivo de recuperación	25
5.2.3. Cinco estilos de juego	26
5.2.3.1. Estilo 1: Estilo ofensivo de posesión	27
5.2.3.2. Estilo 2: Estilo ofensivo directo	27
5.2.3.3. Estilo 3: Estilo defensivo de posesión	28
5.2.3.4. Estilo 4: Estilo equilibrado	29
5.2.3.5. Estilo 5: Estilo defensivo directo	30
5.3. Asignación de estilos de juego	30
6. Conclusiones	34

Bibliografía	36
Anexos	37
A. Distribuciones de cada variable con 3 estilos	37
B. Distribuciones de cada variable con 4 estilos	39
C. Distribuciones de cada variable con 5 estilos	41

Índice de Tablas

3.1.	Equipos Campeonato Nacional 2021	9
3.2.	Matriz de correlación [con DD = Disputas Defensivas]	11
3.3.	Caracterización variables	15
5.1.	Probabilidad de que en un partido cada equipo juegue con cada uno de los estilos	31
5.2.	Probabilidad de que Universidad Católica juegue con cada uno de los estilos en 14 partidos del torneo	32

Índice de Ilustraciones

3.1.	Gráfico disputas defensivas vs disputas defensivas ganadas	10
3.2.	Gráfico disputas defensivas vs porcentaje de disputas defensivas ganadas . . .	10
3.3.	Gráfico presión ofensiva	11
3.4.	Gráfico porcentaje presión ofensiva efectiva	12
3.5.	Gráfico presión ofensiva	12
3.6.	Gráfico porcentaje presión ofensiva efectiva	13
4.1.	Modelo gráfico	17
5.1.	Comportamiento tres estilos	19
5.2.	Comportamiento variables estilo 1	20
5.3.	Comportamiento variables estilo 2	21
5.4.	Comportamiento variables estilo 3	22
5.5.	Comportamiento cuatro estilos	22
5.6.	Comportamiento variables estilo 1	23
5.7.	Comportamiento variables estilo 2	24
5.8.	Comportamiento variables estilo 3	25
5.9.	Comportamiento variables estilo 4	26
5.10.	Comportamiento cinco estilos	26
5.11.	Comportamiento variables estilo 1	27
5.12.	Comportamiento variables estilo 2	28
5.13.	Comportamiento variables estilo 4	29
5.14.	Comportamiento variables estilo 4	29
5.15.	Comportamiento variables estilo 5	30
5.16.	Combinaciones entre estilos más repetidas	32
A.1.	Gráfico comparación parámetro Shots	37
A.2.	Gráfico comparación parámetro Passes	37
A.3.	Gráfico comparación parámetro Crosses	37
A.4.	Gráfico comparación parámetro Defensive challenges won	37
A.5.	Gráfico comparación parámetro Attaking challenges won	38
A.6.	Gráfico comparación parámetro Air challenges won	38
A.7.	Gráfico comparación parámetro Dribbles	38
A.8.	Gráfico comparación parámetro Tackles	38
A.9.	Gráfico comparación parámetro Ball Recoveries	38
A.10.	Gráfico comparación parámetro Building ups	39
B.1.	Gráfico comparación parámetro Shots	39
B.2.	Gráfico comparación parámetro Passes	39
B.3.	Gráfico comparación parámetro Crosses	39
B.4.	Gráfico comparación parámetro Defensive challenges won	39
B.5.	Gráfico comparación parámetro Attaking challenges won	40

B.6.	Gráfico comparación parámetro	Air challenges won	40
B.7.	Gráfico comparación parámetro	Dribbles	40
B.8.	Gráfico comparación parámetro	Tackles	40
B.9.	Gráfico comparación parámetro	Ball Recoveries	40
B.10.	Gráfico comparación parámetro	Building ups	41
C.1.	Gráfico comparación parámetro	Shots	41
C.2.	Gráfico comparación parámetro	Passes	41
C.3.	Gráfico comparación parámetro	Crosses	41
C.4.	Gráfico comparación parámetro	Defensive challenges won	42
C.5.	Gráfico comparación parámetro	Attacking challenges won	42
C.6.	Gráfico comparación parámetro	Air challenges won	42
C.7.	Gráfico comparación parámetro	Dribbles	42
C.8.	Gráfico comparación parámetro	Tackles	43
C.9.	Gráfico comparación parámetro	Ball Recoveries	43
C.10.	Gráfico comparación parámetro	Building ups	43

Capítulo 1

Introducción

1.1. Motivación

En la actualidad los datos son increíblemente importantes, estos permiten a las empresas y organizaciones tomar mejores decisiones, mejorar experiencias de clientes, impulsar la innovación, obtener ventajas competitivas, predecir tendencias y comportamientos futuros, además, de nuevas utilidades que día a día aparecen.

"Sports Analytics" se le llama a la rama que utiliza estadísticas históricas relevantes para brindar una ventaja competitiva a un equipo o individuo. A través de la recopilación y el análisis de estos datos, los análisis deportivos informan a los directores, jugadores, entrenadores y otro personal crucial con el fin de mejorar la toma de decisiones antes, durante y posterior a los eventos deportivos.

El uso de datos en análisis deportivos tiene una larga historia, pero uno de los pioneros en este campo fue Bill James, un escritor y estadístico. James comenzó a analizar datos de béisbol en las décadas de 1970 y 1980, desarrollando nuevas medidas estadísticas que ayudaron a comprender mejor el rendimiento de los jugadores y equipos. El trabajo de James atrajo la atención de los equipos de béisbol y, en la década de 2000, varios equipos de las Grandes Ligas comenzaron a contratar analistas y estadísticos para aplicar análisis de datos avanzados a la evaluación de jugadores, la estrategia de juego y otros aspectos de la gestión del equipo.

Desde entonces, el uso del análisis de datos se ha extendido a otros deportes, como el baloncesto, el fútbol americano y el fútbol. Y se ha convertido en una herramienta cada vez más importante para obtener una ventaja competitiva. Hoy en día, hay muchas personas y organizaciones que trabajan en el campo de la analítica deportiva, incluidos investigadores académicos, analistas de datos y hasta los propios equipos deportivos.

El análisis de fútbol cubre una amplia gama de técnicas y aplicaciones, que incluyen seguimiento de jugadores, análisis de rendimiento, optimización de estrategias de juego y reclutamiento. En particular, en reclutamiento, los clubes profesionales buscan incorporar nuevos jugadores que puedan contribuir a su estrategia de juego para la nueva temporada.

El proceso de reclutamiento en fútbol es el proceso de observar, evaluar y analizar a juga-

dores y equipos de fútbol, con el fin de identificar talentos, fortalezas y debilidades, y tomar decisiones informadas sobre posibles fichajes y estrategias de juego. Este, puede ser llevado a cabo por los clubes de fútbol de manera interna, o pueden contratar a empresas especializadas en este ámbito.

Actualmente, traen jugadores que triunfaron en otros clubes, sin tener necesariamente en cuenta la diferencia entre los estilos de juego de los respectivos equipos. Esto plantea la tarea de integrarlos en la dinámica del equipo y adaptarlos al plan de juego. Es importante para los equipos que estos nuevos fichajes logren adaptarse lo más rápido posible, para que rindan al mismo nivel que en sus clubes anteriores, o mejor aún, que obtengan rendimientos superiores.

El proceso, aunque es esencial para la toma de decisiones informadas por parte de los clubes, tiende a ser un proceso lleno de sesgos. Esto se debe a que la evaluación de los jugadores y equipos puede estar influenciada por las percepciones personales y las preferencias del reclutador, lo que puede llevar a decisiones equivocadas en cuanto a la selección.

Como relatan Tom L. G. Bergkamp (2022) en el “European Journal of Sports Science“ si bien, los reclutadores evalúan los atributos de los jugadores de manera estructurada, en última instancia, basan su predicción en una integración intuitiva de diferentes atributos de rendimiento. Al depender de la observación subjetiva, los reclutadores pueden tomar decisiones a partir de prejuicios sobre ciertos jugadores, de sus características físicas o simplemente pasar por alto información crucial que podría ser proporcionada por los datos.

Si el jugador escogido no cumple con las expectativas una vez que se une al club, puede resultar en una pérdida financiera significativa, especialmente si se trata de un contrato a largo plazo. El salario de un jugador, así como otros costos asociados con su contratación, pueden representar una gran inversión para un club de fútbol. Además, si el jugador no rinde al nivel esperado, podría llevar al equipo a perder partidos, lo que afectaría su posición en la tabla y, en última instancia, su capacidad para generar ingresos. Por lo tanto, es importante que los clubes de fútbol tomen decisiones de contratación informadas y consideren cuidadosamente el desempeño pasado y el potencial futuro de un jugador antes de invertir en él.

Podemos ver muchos ejemplos de esta situación a lo largo de la historia del fútbol profesional: como por ejemplo, el de Fernando Torres relatado por Moore (2020), quien pasó del Liverpool al Chelsea en 2011 por una tarifa récord de transferencia (58 millones de euros). Torres había sido uno de los delanteros más prolíficos de la Premier League mientras jugaba para el Liverpool, pero su desempeño cayó drásticamente después de mudarse al Chelsea, donde tuvo dificultades para adaptarse al estilo de juego y la estructura del equipo. A pesar de que ganó la Liga de Campeones y la Europa League con el Chelsea, Torres nunca logró igualar su éxito en el Liverpool y finalmente fue vendido al AC Milán en 2014.

Los datos pueden ayudar a identificar patrones y tendencias en el rendimiento de los jugadores, lo que puede ayudar a predecir su éxito en el futuro. Al incorporar datos objetivos en el proceso de reclutamiento, los equipos pueden tomar decisiones más informadas y precisas sobre la selección de jugadores, lo que puede tener un impacto significativo en su éxito deportivo y financiero.

1.2. Objetivo y Metodología

Es por esto todo lo mencionado anteriormente, que el objetivo de esta investigación es detectar estilos de juego en equipos de fútbol profesional. Para identificar cuáles son similares, y mejorar la adaptación del nuevo jugador en un nuevo potencial club. Todo esto, mediante técnicas estadísticas y datos.

Para lograr esto, se propone un método de clasificación no supervisado. En particular, un modelo de inferencia bayesiana, para proporcionar un marco probabilístico para las relaciones entre las métricas de rendimiento de cada equipo y los estilos de juego. Incorpora conocimientos previos sobre los datos y ajusta el modelo en consecuencia.

El uso de un método de clasificación para detectar estilos de juego puede ayudar a mejorar la objetividad, la precisión y la utilidad de la información obtenida, lo que puede ser beneficioso para el análisis, la toma de decisiones y la estrategia en el fútbol. Además, al conocer los diferentes estilos de juego de sus oponentes, los equipos pueden adaptar su estrategia y tácticas de juego para contrarrestar las fortalezas y debilidades de sus rivales.

Este proyecto utiliza datos del campeonato nacional 2021 de la primera división de Chile. Estos, se adquirieron a través de InStat, empresa de análisis de rendimiento deportivo. Para cada una de los equipos se consideran todas las fechas de la temporada 2021, con variables que describen de forma agregada lo que sucedió en la cancha a nivel partido y por equipos, tales como: disputas defensivas, disputas ofensivas, pases, tiros al arco, regates, etc.

El modelo bayesiano propuesto estima la distribución de probabilidad del parámetro “estilo de juego”, dados los datos. Esto nos permite, a través de un análisis de los resultados, declarar ciertos “estilos de juego” y qué equipos pertenecen a cada uno con cierta probabilidad. Posteriormente, se realizará un análisis de sensibilidad para ajustar mejor el modelo y definir parámetros tales como cuántos clusters se van a obtener, cuántas variables de rendimiento se deben considerar en los datos y, más importante, cuáles específicamente.

Cabe destacar que para uso de esta investigación se definirá “estilo de juego” como la combinación de patrones de rendimiento en la cancha de un equipo. No se considerará la formación táctica ni la filosofía del entrenador.

1.3. Contribución

Esta investigación corresponde a un primer acercamiento a la clasificación de estilos de juego de equipos a través de inferencia Bayesiana. El modelo podría potencialmente ser extrapolado a diferentes ligas, de muchos países e incluso en otros deportes que sean en equipo y se puedan medir métricas de performance.

Además, es importante recalcar la ausencia de investigaciones similares en el contexto chileno. Esta combinación única de enfoques contribuye al conocimiento del fútbol chileno y proporciona información valiosa para el desarrollo de futuras investigaciones.

Capítulo 2

Revisión de Literatura

En esta sección, se busca explicar los conceptos fundamentales que se relacionan con el tema de investigación. Además, se busca identificar y analizar las investigaciones previas relevantes para la investigación en cuestión, lo que ayuda a contextualizar el problema a abordar, y a identificar vacíos o áreas de mejora en la literatura existente.

2.1. Clasificación de estilos de juego

Estilo de juego según Adam Hewitt (2016), es un término comúnmente utilizado por entrenadores, analistas de rendimiento y medios de comunicación para referirse a patrones de juego en deportes de equipo. Un método de clasificación permite identificar estos patrones y características comunes en el juego, lo que puede ser útil para la identificación, análisis y la comparación de diferentes estilos de juego.

Existe evidencia del uso de clasificación como método de análisis dentro del "Sports Analytics", y más específicamente, del fútbol. Gregory (2019) proporciona un método para identificar y clasificar carreras sin balón en posesión en grupos similares para permitir un análisis más generalizable. El objetivo era crear un vocabulario de tipos de ejecución que se pueda usar para describir o analizar mejor las ejecuciones específicas dentro de un partido, y que se pueda consultar más fácilmente que los datos de seguimiento sin procesar.

En particular, la literatura académica cuenta con numerosos estudios enfocados en la identificación de estilos de juego en el fútbol. Podemos observar el trabajo de Jeroen Clijmans (2023), en el que se capturan los patrones secuenciales del estilo de un equipo, modelando el comportamiento observado como una cadena de Markov en tiempo discreto. En segundo lugar, caracterizan el estilo ofensivo de los equipos en una serie de rasgos. Aplican una combinación de técnicas analíticas y comprobación probabilística del modelo para razonar sobre el esquema de un equipo, con el fin de extraer valores para estas características. Todo esto, usando datos de flujo de eventos de la Premier League inglesa 2019-2020. El modelo resultante permitía encontrar equipos con estilos de juego similares, encontrar ineficiencias en el estilo de juego, y realizar un análisis en profundidad. Cabe destacar, que este estudio no logra entregar estilos de juego generales, sino que similitudes entre equipos de una liga.

Muchos de los estudios que buscan identificar los estilos de juego se han enfocado, en gran medida, en estilos ofensivos y defensivos. Estos dos aspectos son considerados como los más

importantes en el juego, ya que determinan la forma en que el equipo ataca y defiende.

Un ejemplo de esto es el trabajo de Drezner, Lamas, Barrera, y Dantas (2020), quienes aplican un conjunto de variables relacionadas con la circulación del balón para evaluar nueve partidos de los finalistas de la UEFA Champions League temporada 2008-2009, Barcelona y Manchester United. Se realizó una evaluación a través de una identificación automática de los eventos del juego usando un software de máquina de estado finito, que buscaba clases particulares de secuencias de codificación en los datos de clases de circulación de balón adquiridos manualmente, a partir de secuencias de vídeo. En general, el enfoque utilizado en este trabajo fue eficaz para detectar diferencias entre los dos equipos, principalmente en relación con el estilo general de circulación del balón. Sin embargo, no se declaran categorías generales.

Más específicamente, Fernández Navarro (2019) utiliza análisis factorial para detectar estilos de juego en fútbol utilizando datos de la Premier League inglesa y la La Liga española. El análisis factorial es una técnica estadística que se utiliza para analizar la estructura de interrelaciones entre varias variables observadas. Reduce la dimensionalidad de un conjunto de datos, identificando patrones subyacentes y agrupando las variables en grupos llamados factores.

Plakias S (2023) realizó una revisión crítica de 40 artículos de investigación que buscaban estilos de juego. El análisis descriptivo y temático encontró que los objetivos de los artículos identificados se pueden clasificar en tres categorías principales (reconocimiento, efectividad de los estilos de juego y las variables contextuales que los afectan). El documento concluyó que: hasta ahora el análisis factorial parece ser la mejor técnica usada pero requiere atención a la interpretación de los componentes; la inteligencia artificial abre nuevos horizontes, y finalmente, que existe una necesidad de más investigación sobre la efectividad de los diferentes estilos de juego.

Las conclusiones fueron que a pesar de la importancia de la información que ofrecen estos estudios, no pueden cubrir la complejidad del juego de fútbol, que requiere la conexión de diferentes indicadores de rendimiento para reconocer los estilos de juego. Por lo tanto, se requiere más investigación sobre los estilos de juego que se centre en la aplicación real de los hallazgos en los campos de fútbol.

Estas investigaciones han utilizado diversas metodologías para analizar las variables que pueden influir en la forma en que los equipos juegan. Sin embargo, cabe destacar que la búsqueda de estilos de juego en el fútbol es un tema complejo debido a la multitud de factores que pueden influir en el rendimiento de los equipos, como las tácticas, estrategias, habilidades individuales de los jugadores, entre otros.

En particular, casi todos los estudios se han realizado con datos de países europeos o asiáticos. El fútbol en Sudamérica se distingue de manera notable del fútbol en Europa y Asia, tanto en términos tácticos como en su estilo de juego. Esta investigación es un primer acercamiento con datos de fútbol sudamericano.

2.2. Inferencia Bayesiana

La Inferencia Bayesiana es un método que se usa para poder estimar algún elemento probabilístico que se desconozca en exactitud, observando la evidencia. Está basada en la interpretación subjetiva de la probabilidad y tiene como punto central el Teorema de Bayes.

Es posible usarla como método de clasificación debido a que proporciona un marco probabilístico para modelar las relaciones entre los puntos de datos y las agrupaciones. Permite estimar la distribución de probabilidad de los parámetros de un modelo, dados los datos, y usar estas estimaciones para hacer predicciones sobre nuevos puntos de datos.

Podemos notar, que se ha utilizado modelos bayesianos como método de clasificación no supervisado, un ejemplo de esto es el estudio de Jurek y Zakrzewska (2008). Consideran la aplicación del modelo Naive Bayes, para la evaluación del riesgo relacionado con los seguros de vida de los clientes. Los clientes se clasifican en grupos de diferentes niveles de riesgo de seguro.

Han surgido bastantes acercamientos al Sports Analytics a través de Inferencia Bayesiana en los últimos años, sobre todo en béisbol. Deshpande y Wyner (2017) presentan un modelo jerárquico bayesiano para estimar la probabilidad de que cada árbitro cante un strike, ajustando los participantes del campo, la ubicación del campo y la información contextual como el conteo. Luego, se puede traducir estos efectos estimados en carreras promedio ahorradas durante una temporada.

Maddox, Sides, y Harvill (2022) en cambio, proponen dos nuevos métodos bayesianos para estimar y predecir las probabilidades de victoria del equipo local en el juego del baloncesto universitario masculino de la División I de la NCAA. El primer método tiene un ajuste previo que se ajusta en función del diferencial de plomo y el tiempo transcurrido. El segundo es una versión ajustada del primero, donde el ajuste es una combinación lineal del estimador bayesiano con una probabilidad de ganar antes del juego ponderada en el tiempo.

En fútbol, solamente se pueden encontrar modelos bayesianos con el fin de predecir resultados o para inferir habilidad, como lo hizo Whitaker, Silva, Edwards, y Kosmidis (2017). Quienes se propusieron determinar la habilidad de un jugador de fútbol para un evento dado, por ejemplo, anotar un gol. Implementaron un modelo de Poisson para capturar ocurrencias de tipos de eventos, a partir de los cuales inferir las habilidades del jugador. Este enfoque también permite la visualización de diferencias entre jugadores, por una habilidad específica. Todo esto con datos de la Premier League inglesa, capturando las habilidades de los jugadores durante la temporada 2013-2014, antes de usar resultado del modelo jerárquico para predecir si se marcarán más o menos de 2,5 goles en un juego dado en la temporada 2014-2015.

Baio y Blangiardo (2010) proponen un modelo jerárquico Bayesiano para abordar las características que hacen que un equipo pierda o gane un partido y para predecir el marcador de un partido en particular, en datos sobre el campeonato de la Serie A italiana 1991-1992.

Karlis y Ntzoufras (2008) también proponen un modelo bayesiano con el objetivo de predecir resultados, sin embargo, a diferencia del anterior, este se basaba en la diferencia de goles

usando la llamada distribución de Skellam. Se utilizan datos de la Premier league inglesa pero esta vez, para la temporada 2006-2007. Es por todo lo anterior, que si bien no se ha utilizado la clasificación a través de inferencia bayesiana para estudiar fenómenos en fútbol, es una técnica que tiene un gran potencial para ser aplicada.

En resumen, esta investigación se destaca como un estudio innovador y prometedor en el campo, ya que aborda un tema de gran relevancia sin haber sido explorado previamente desde la perspectiva de la Inferencia Bayesiana. Además, es notable la ausencia de enfoques similares en el contexto del fútbol latinoamericano, específicamente en el contexto chileno. Esta combinación única de enfoques metodológicos y geográficos agrega un valor significativo a la investigación, proporcionando una comprensión más profunda y precisa de los patrones y dinámicas del fútbol en una región. Por lo tanto, este estudio no solo contribuye al cuerpo existente de conocimiento, sino que también sienta las bases para futuras investigaciones y posibles aplicaciones prácticas en el ámbito del fútbol latinoamericano y su desarrollo estratégico.

Capítulo 3

Datos

En la siguiente sección se explica el origen de los datos y análisis exploratorio de datos para tener mayor claridad de la estructura y la calidad de con qué se cuenta.

3.1. Proveedor

Los datos utilizados en esta investigación provienen de InStat, empresa proveedora de datos deportivos que se especializa en fútbol y fueron provistos por el ISCI, Instituto de Sistemas Complejos de Ingeniería a través de un fondo concursable. Instat, utiliza tecnología avanzada para recopilar y analizar una gran cantidad de información en tiempo real durante los partidos, incluyendo estadísticas detalladas sobre jugadores, equipos y jugadas.

En particular, en esta investigación se utilizan datos del torneo nacional de Chile 2021 (Campeonato PlanVital 2021). Para cada una de las fechas se consideran 117 variables que describen de forma agregada lo que sucedió en la cancha a nivel partido y por equipos (dos observaciones por cada partido, una para cada equipo) tales como: disputas defensivas, disputas ofensivas, pases, tiros al arco, regates, etc. En total se cuenta con 17 equipos, los cuáles pueden ser observados en la tabla 3.1.

3.2. EDA

Con tal cantidad de datos, es indispensable reducir la dimensionalidad. Realizar un buen análisis exploratorio de los datos con los que se cuenta es clave, porque nos permite entender de mejor forma cómo están construidas las variables que nos servirán para detectar los estilos de juego. Para empezar, se realiza un tratamiento de los datos, es decir, eliminar las filas con datos faltantes, mal ingresados o erróneos (no se puede tener un jugador que juegue 150 min en un partido por ejemplo).

Luego, se procede a estudiar los datos existentes, sus promedios y varianzas. Esto, debido a que varían mucho en rango (por ejemplo: 7.000 seg de posesión del balón en seg vs 14 pases efectivos). El objetivo principal de esta técnica es buscar formas similares en los datos.

También, es importante estudiar si algunas variables están correlacionadas. Esto, con el

Tabla 3.1: Equipos Campeonato Nacional 2021

ID	Team	Sigla
1	Club de Deportes Provincial Curicó Unido	CUR
2	Club de Deportes Melipilla	MEL
3	Club Deportivo O'Higgins	OHI
4	Everton de Viña del Mar	EVE
5	Deportes Unión La Calera	ULC
6	Club Social y Deportivo Colo-Colo	CC
7	Club de Deportes Cobresal	COB
8	Huachipato FC	HUA
9	Club de Deportes Antofagasta	ANT
10	Club Deportivo Palestino	PAL
11	Unión Española	UE
12	Club de Deportes Santiago Wanderers	SW
13	Club Deportivo Universidad Católica	UC
14	Club Deportivo Ñublense	ÑUB
15	Audax Italiano La Florida	AUD
16	Club de Deportes La Serena	SER
17	Club Universidad de Chile	UCH

fin de identificar y eliminar datos relacionados linealmente y que nuestro modelo futuro sea más eficiente. El resultado es que existen bastantes variables altamente correlacionadas tales como: “Pases” vs “Pase efectivos” en un 98 %, “Porcentaje de posesión del balón” vs “Posesión del balón en seg” en un 97 %, etc. Lo que genera la difícil decisión de elegir qué variables son las que serán utilizadas en la investigación.

La discusión generada a partir de esto es de qué define realmente el estilo de juego de un equipo, lo que se intenta hacer o lo que efectivamente se logra, por lo que podemos notar en los datos no debería haber mucha diferencia (altamente correlacionados), sin embargo, la intuición llama a decir que lo que efectivamente se logra es lo observable y definirá cómo es el estilo de juego.

Se puede observar que dentro de la base se contaban con varias variables ligadas al mismo tema pero miradas desde enfoques distintos, por ejemplo: “disputas defensivas”, “disputas defensivas ganadas” y “porcentaje de disputas defensivas ganadas”. Mirando las “disputas defensivas” vs las “disputas defensivas ganadas” en la imagen 3.1 se ve una relación lineal positiva, lo que indica que se comportan de la misma forma y estarían correlacionadas.

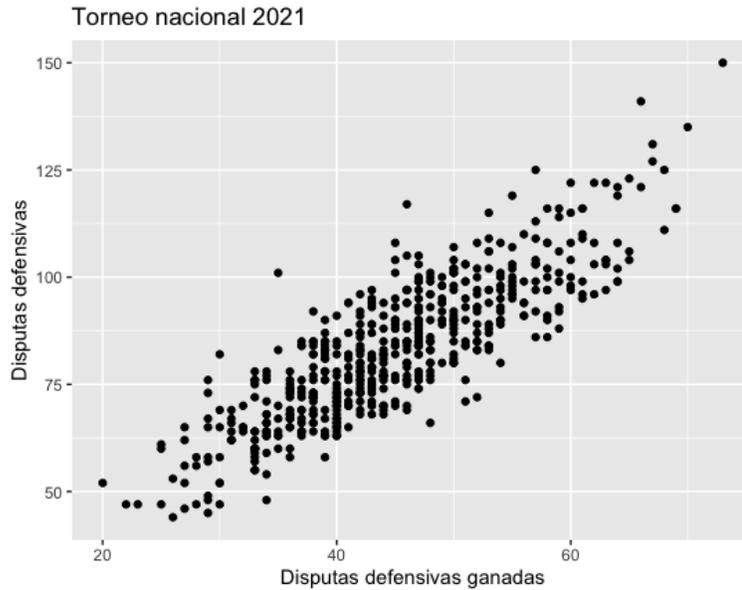


Figura 3.1: Gráfico disputas defensivas vs disputas defensivas ganadas

A mayor cantidad de disputas defensivas, mayor también fueron las que efectivamente fueron ganadas. No ocurrió lo mismo con el porcentaje de disputas defensivas ganadas como se puede observar en la imagen 3.2.

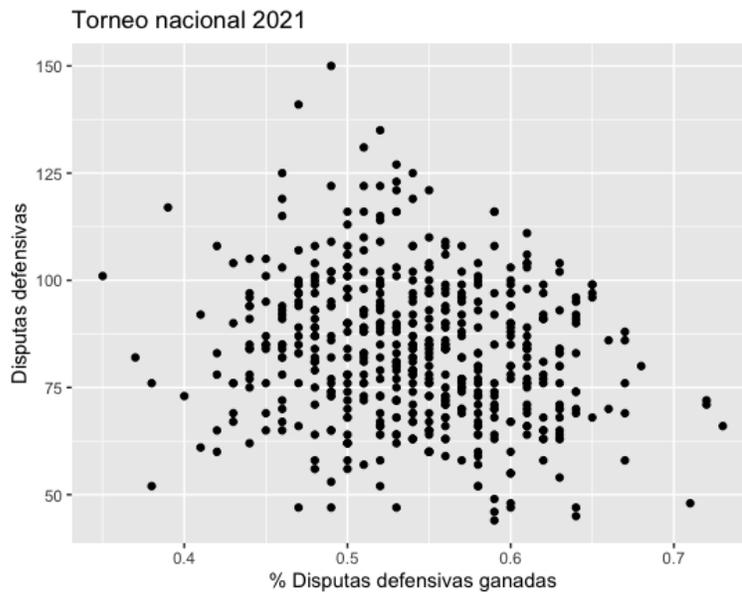


Figura 3.2: Gráfico disputas defensivas vs porcentaje de disputas defensivas ganadas

Al observar la matriz de correlación de la tabla 3.2, se concluyó que no se podía considerar ambas, disputas defensivas y disputas defensivas ganadas, esto, permite eliminar la correlación. Este será el criterio usado para todas las variables en las que ocurría la misma situación.

Además de lo mencionado anteriormente, no se considerará variables que se enfocan en el rendimiento individual, como goles, penales u otros indicadores similares, para dar prioridad

Tabla 3.2: Matriz de correlación [con DD = Disputas Defensivas]

	DD	DD ganadas	Porcentaje de DD ganadas
DD	1	0.85	-0.17
DD ganadas	0.85	1	0.35
Porcentaje de DD ganadas	-0.17	0.35	1

a la evaluación colectiva del equipo.

Realizando una observación un poco más profunda de los datos se puede notar desde ya que ciertos equipos son más ofensivos, y otros son más defensivos.

Para analizar la presión ofensiva de todos los equipos del torneo nacional 2021, se muestra la imagen 3.3, con los equipos en orden de cómo terminaron en la tabla de posiciones final (el campeón a la izquierda y los que descienden a la derecha). Ahora, observando la imagen 3.4 vemos cuanto porcentaje de esa presión fue efectiva.

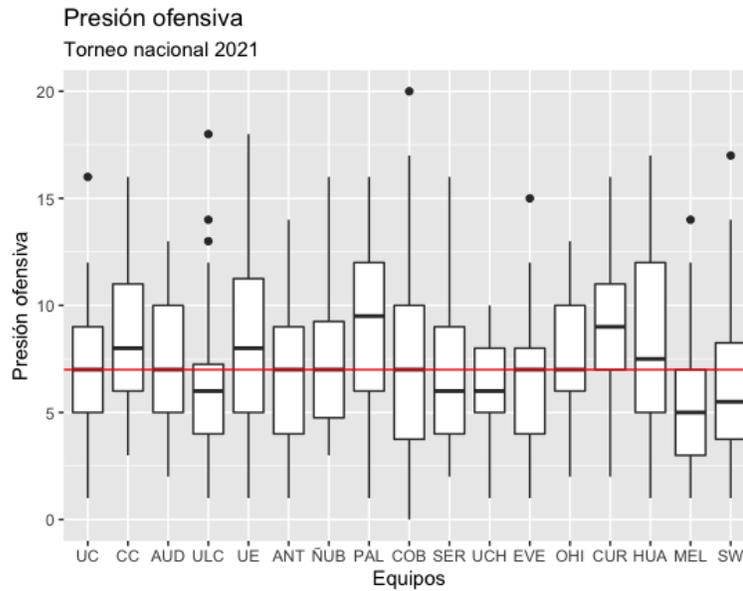


Figura 3.3: Gráfico presión ofensiva

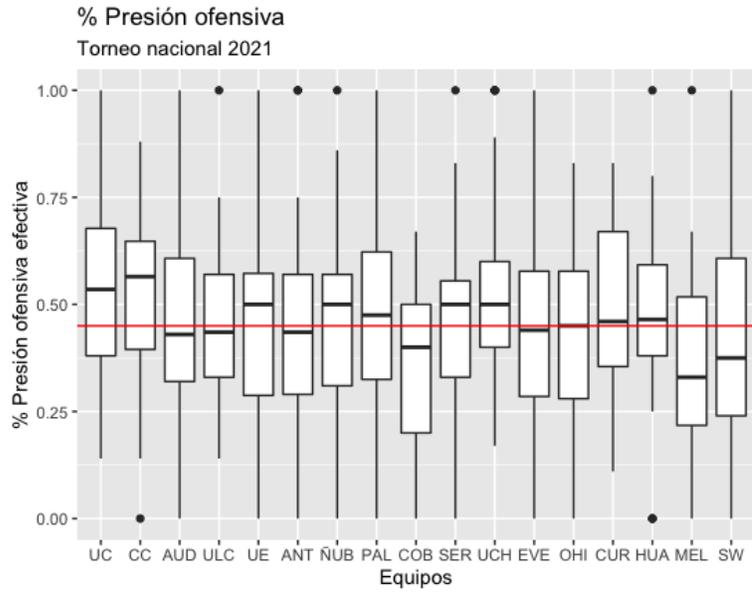


Figura 3.4: Gráfico porcentaje presión ofensiva efectiva

Podemos notar que, si bien la presión que se intentó, varía mucho según la posición en la que terminaron. Equipos como Colo Colo y Unión Española muestran alta presión ofensiva pero además, alta precisión en esta.

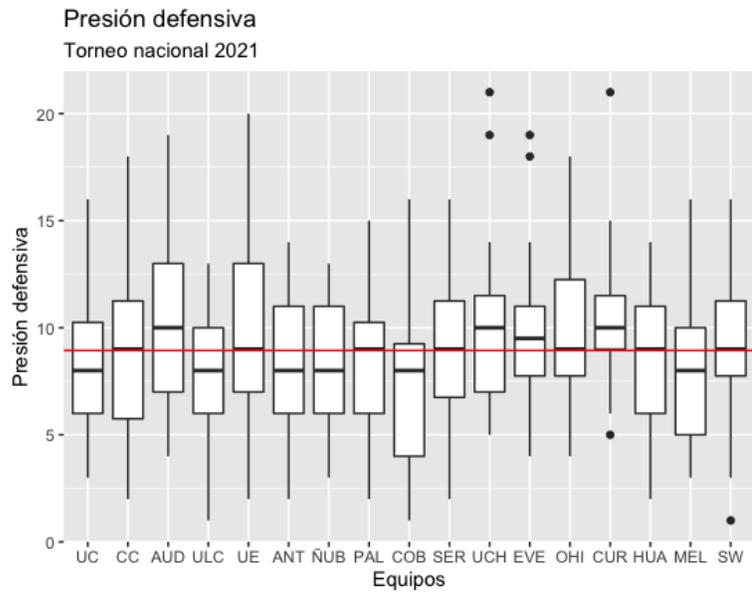


Figura 3.5: Gráfico presión ofensiva

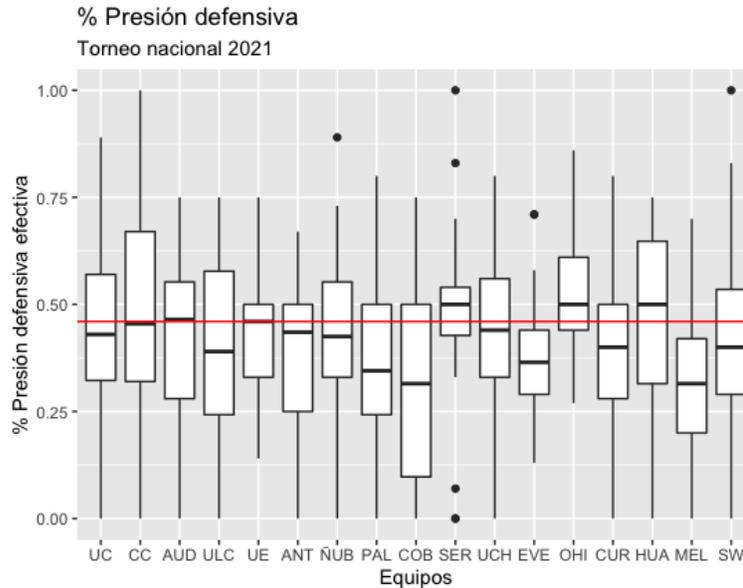


Figura 3.6: Gráfico porcentaje presión ofensiva efectiva

En cambio, si miramos la presión defensiva de todos los equipos del torneo nacional 2021 en la imagen 3.5, y luego lo comparamos con el porcentaje que fue realmente efectivo en la imagen 3.6, notamos que equipos como O’Higgins, La Serena y Huachipato muestran una alta presión defensiva y alta precisión en esta. Ya que se está estudiando “estilos de juego” este entendimiento es muy importante al momento de definir qué estilos de juego se considerarán y que características tiene cada uno.

3.3. Selección de variables

Para el siguiente modelo bayesiano, es crucial considerar cuidadosamente el número de variables que se utilizan para caracterizar los estilos post resultados. Una gran cantidad de variables puede complicar significativamente el análisis y la interpretación de los resultados.

Cuando el número de variables es demasiado grande, el modelo puede volverse difícil de manejar, ya que se requiere una mayor capacidad de recursos para estimar las distribuciones de probabilidad conjuntas de todas las variables. Es importante tener en cuenta que cada variable adicional introduce una mayor incertidumbre en el modelo y puede requerir una cantidad significativa de datos para estimar sus parámetros con precisión. Además, una gran cantidad de variables puede llevar a problemas de sobreajuste, donde el modelo se adapta demasiado a los datos de entrenamiento y no generaliza bien a nuevos datos.

Limitar el número de variables garantiza una caracterización más efectiva de los estilos, facilitar el análisis e interpretación de los resultados y evitar problemas de sobreajuste y complejidad computacional.

La selección de variables descriptivas es fundamental para capturar de manera efectiva lo que el equipo realmente jugó en el campo. Al hacerlo, se logra una caracterización más

precisa y representativa del desempeño del equipo, lo que a su vez facilita la interpretación y el análisis del estilo.

Luego de un exhaustivo análisis de las 110 variables con las que se contaba en la base, se seleccionan 10 que permiten hablar de forma general como se comportó el equipo en la cancha:

- Shots (Tiros): Representa el número de intentos de tiros al arco realizados un equipo durante un partido. Es un indicador de las oportunidades de gol generadas.
- Passes (Pases): Indica el número de pases exitosos realizados por el equipo durante el partido.
- Crosses (Centros): Refiere a los tiros de balón desde los costados hacia el área rival en busca de un rematador.
- Defensive challenges (Disputadas defensivas ganadas): Muestra la cantidad de duelos defensivos en los que el jugador salió victorioso.
- Attacking challenges won (Disputas ofensivas ganadas): Representa los duelos ofensivos en los que el jugador salió victorioso.
- Air challenges won (Disputas aéreas ganados): Indica el número de duelos aéreos ganados por el jugador, ya sea en defensa o en ataque.
- Dribbles (Regates): Hace referencia a la cantidad de veces que el jugador ha superado a un adversario mediante habilidades técnicas individuales con el balón.
- Tackles (Barridas): Representa la cantidad de entradas o intervenciones defensivas exitosas realizadas por el jugador para recuperar el balón.
- Ball recoveries (Recuperaciones del balón): Indica la cantidad de veces que un jugador del equipo ha recuperado el balón.
- Building ups (Construcción de juego): Refiere a las acciones del jugador que contribuyen a la creación de jugadas ofensivas o al inicio de una transición controlada desde la defensa hacia el ataque.

Para facilitar el análisis posterior, se puede observar la caracterización y los valores que toman las variables seleccionadas para todos los equipos en el Torneo Nacional 2021 en la tabla 3.3.

Para concluir, se logra un buen entendimiento de los datos con los que se trabajará próximamente y se disminuye significativamente la dimensionalidad. A partir de los criterios usados anteriormente se seleccionan 10 variables a considerar en el modelo próximamente descrito.

Tabla 3.3: Caracterización variables

ID	Variable	Promedio	Std	Mínimo	Máximo
Variable 1	Shots	11.51	4.56	2.0	29.0
Variable 2	Passes	424.43	85.11	175.0	758.0
Variable 3	Crosses	13.22	6.48	1.0	43.0
Variable 4	Defensive challenges won	44.83	9.46	20.0	73.0
Variable 5	Attacking challenges won	38.48	9.96	14.0	77.0
Variable 6	Air challenges won	20.82	8.0	1.0	49.0
Variable 7	Dribbles	26.37	8.61	7.0	68.0
Variable 8	Tackles	34.19	9.00	10.0	78.0
Variable 9	Ball recoveries	55.12	7.86	33.0	75.0
Variable 10	Building ups	29.47	8.43	9.0	54.0

Capítulo 4

Modelo

En la siguiente sección, se presentará el modelo generativo de datos. Este crucial componente de la investigación nos permitirá comprender cómo se relacionan y comportan los datos utilizados.

4.1. Modelo probabilista

Siendo S el número de observaciones de n variables de rendimiento de partidos para múltiples equipos en una liga. Sea E el conjunto de equipos a considerar y P el conjunto de partidos totales.

Supondremos que al momento de enfrentar un partido, cada equipo elige un estilo de juego. Los estilos de juego determinan el comportamiento de los equipos durante los partido, y por lo tanto correlacionan con los estadísticos descriptivos como pases, regates, etc, que resumen su transcurso. Existen M estilos de juego, y el estilo de juego m se caracteriza por un vector representativo $r^m \in \mathbb{R}^n$. Supondremos que

$$r^m \sim \mathcal{N}(\mu_0, \text{diag}(\tau_0)).$$

$\mu_0 \in \mathbb{R}^n$, vector de longitud n que representa la media a priori para el vector de parámetros de estilo de juego para cada variable m . Adicionalmente,

$$\mu_0 \sim \mathcal{N}(0, 1).$$

$\tau_0 \in \mathbb{R}_+^n$ es un vector de longitud n que representa la precisión a priori para el vector de parámetros de estilo de juego para cada variable m .

$$\tau_0 \sim \text{Gamma}(2, 0.1).$$

Definimos $R \in \mathbb{R}^{M \times n}$ como una matriz que representa los parámetros de estilo de juego para cada uno de los M estilos de juego y n variables.

Un equipo $e \in E$, en un partido cualquiera $p \in P$, escoge jugar con el estilo $m \in M$ con probabilidad Q_m^e . Este es un vector de longitud M que representa las probabilidades de que cada equipo en la muestra juegue con cada uno de los M estilos de juego. Supondremos que

$$Q_m^e : m \in M, e \in E \sim \text{Dirichlet}(1_M).$$

Hay un vector de probabilidades Q^e para cada uno de los e equipos en la muestra.

Siendo X_p^e la matriz de estadísticos que describe las observaciones de n variables de rendimiento para todas las observaciones S de la muestra. Y siendo J_p^e el estilo de juego escogido por el equipo e durante el partido p .

La descripción del modelo es

- $X_p^e | J_p^e \sim \mathcal{N}(r^{J_p^e}, \text{diag}(\tau_0))$
- $J_p^e \sim \text{Categorical}(Q^e)$
- $Q^e \sim \text{Dirichlet}(1_M)$
- $r^m \sim \mathcal{N}(\mu_0, \text{diag}(\tau_0))$
- $\tau_0 \sim \text{Gamma}(2, 0.1)$
- $\mu_0 \sim \mathcal{N}(0, 1)$

En la imagen 4.1 se puede observar el modelo representado en forma gráfica, con el objetivo de entender mejor como se relacionan las variables entre ellas.

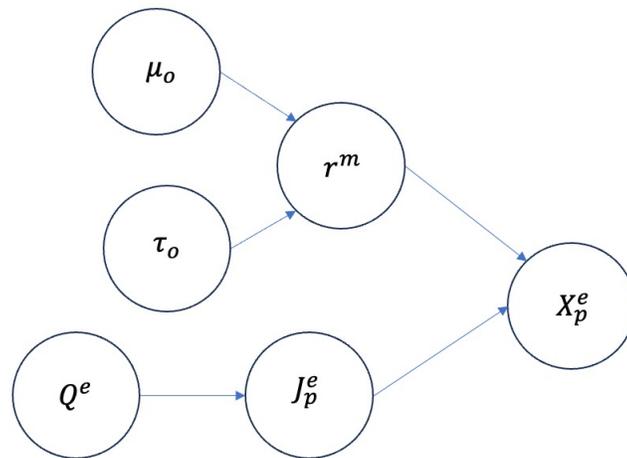


Figura 4.1: Modelo gráfico

Capítulo 5

Resultados

5.1. Monte Carlo Markov Chain

Para resolver el modelo, se utilizó el lenguaje de programación Stan, el cual fue implementado en Python utilizando la interfaz CmdStan. En particular, se utilizó el algoritmo de Monte Carlo Markov Chain (MCMC) Hamiltoniano (HMC) para la inferencia bayesiana.

El algoritmo Hamiltoniano (HMC) es un método que se utiliza para muestrear distribuciones de probabilidad complejas. En general, HMC es un algoritmo muy poderoso y efectivo para la inferencia bayesiana, ya que puede generar muestras más eficientemente que otros algoritmos de MCMC. Una de las principales ventajas de HMC es su capacidad para explorar la distribución de probabilidad de manera más eficiente que los métodos de MCMC que solo utilizan propuestas aleatorias. Además, HMC puede manejar mejor las correlaciones entre las variables.

Para resolver el modelo en Stan, se comenzó definiendo los datos requeridos para la inferencia, incluyendo el número de variables, equipos y partidos, así como las observaciones y estilos de juego iniciales. Luego se especificaron los parámetros que necesitaban ser inferidos, incluyendo la media y la precisión de los parámetros del estilo de juego, así como los parámetros de estilo de juego. Después, se formuló el modelo, que incluyó las distribuciones de probabilidad previas para los parámetros del estilo de juego y las probabilidades del estilo de juego, así como la verosimilitud de los datos. Finalmente, se ajusta el modelo a los datos utilizando el algoritmo de HMC. Durante este proceso, Stan estima la distribución posterior de los parámetros.

5.2. Clasificación estilos de juego

Debido a que el modelo requiere que se le entregue el número de estilos a clasificar, el modelo fue corrido para las 10 variables seleccionadas con número de estilos igual a 2, 3, 4, 5 y 6. De esta forma es posible comparar los resultados y tomar una decisión más informada. Para efectos de este documento se mostrarán los resultados para número de estilos igual a 3, 4 y 5 ya que 2 estilos simplifica mucho la complejidad del juego y 6 estilos genera una sobre segmentación.

5.2.1. Tres estilos de juego

Se procede a describir el comportamiento de las variables para cada uno de los 3 estilos detectados. En la sección A del Anexo se pueden ver las distribuciones considerando más iteraciones en detalle. En la imagen 5.1 se puede observar las gráficas para cada estilo en comparación.

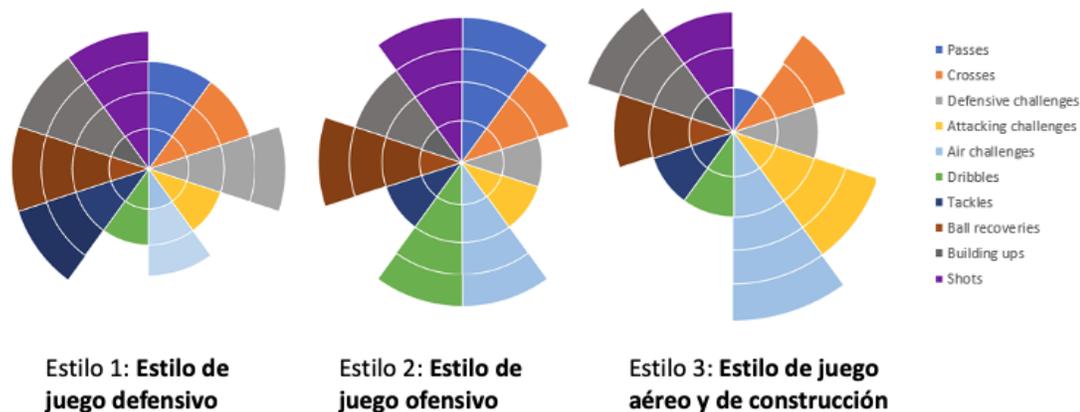


Figura 5.1: Comportamiento tres estilos

5.2.1.1. Estilo 1: Estilo de juego defensivo

El estilo de juego se centraría en una defensa sólida, con una alta tendencia a ganar disputas defensivas y recuperar el balón. Este estilo de juego tiene una tendencia alta en tiros al arco. En cuanto a los pases, hay una baja tendencia, lo que podría sugerir que el equipo no se basa tanto en el juego de pases. Los centros tienen una tendencia promedio, lo que indicaría que el equipo utiliza esta estrategia ocasionalmente, pero no es un aspecto central de su juego.

En las disputas defensivas, el equipo tiene una alta tendencia a ganarlas. En contraste, en las disputas ofensivas que tienen una tendencia levemente baja. Las disputas aéreas tienen una muy baja tendencia, lo que indicaría que el equipo no suele utilizar mucho el juego aéreo.

En cuanto a los regates, el equipo tiene una muy alta tendencia. En las barridas, el equipo muestra una alta tendencia. Esto indicaría que son agresivos en la defensa y no temen arriesgarse para despejar el balón o evitar que el equipo contrario avance. Las recuperaciones de balón tienen una alta tendencia. Y finalmente, los building ups tienen una tendencia levemente baja.

En resumen, este estilo de juego se caracterizaría por enfocarse en una defensa sólida. Prefiriendo la recuperación de balón y los regates. En la imagen 5.2 se puede observar el resumen de la caracterización según tendencias muy alta, alta, promedio, baja y muy baja.

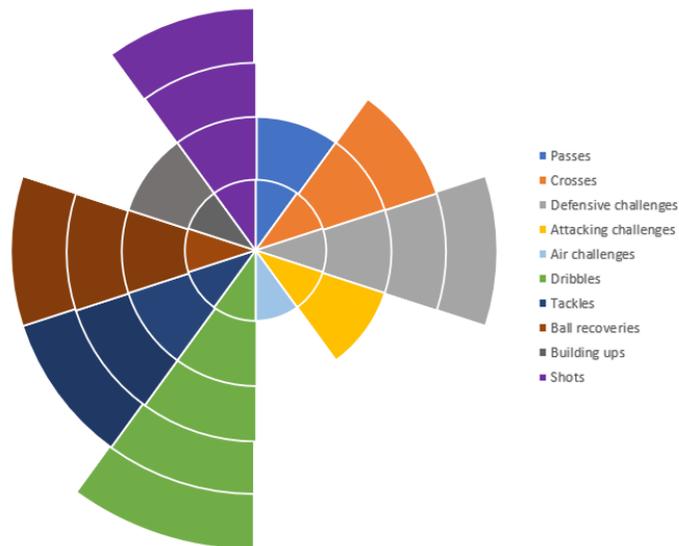


Figura 5.2: Comportamiento variables estilo 1

5.2.1.2. Estilo 2: Estilo de juego ofensivo

Este estilo de juego se centra en la ofensiva, con una alta tendencia tanto en tiros como en pases, lo cual indicaría que están constantemente buscando oportunidades de gol. Además, se observa una alta tendencia en regates. En cuanto a los centros, se observa una tendencia promedio.

En términos defensivos, el equipo muestra una baja tendencia en las disputas defensivas ganadas y las barridas. En cuanto a las disputas ofensivas ganadas, también se observa una tendencia baja. Las disputas aéreas ganadas muestran una alta tendencia, lo que sugiere que el equipo es efectivo en el juego aéreo, ya sea en defensa o en ataque.

En términos de recuperaciones de balón, se observa una levemente alta tendencia. Esto indica que el equipo tiene una habilidad decente para recuperar el balón después de perderlo, lo que puede ser importante para mantener la presión sobre el oponente y generar nuevas oportunidades de ataque.

Finalmente, los building ups muestran una tendencia promedio. Esto sugiere que el equipo no se enfoca exclusivamente en la construcción lenta y elaborada del juego.

En general, este estilo de juego se caracteriza por un enfoque ofensivo con énfasis en los tiros, pases, regates y disputas aéreas ganadas. Aunque muestran una menor tendencia en las disputas defensivas ganadas y las barridas, compensan con una levemente alta tendencia en recuperaciones de balón. En la imagen 5.3 se puede observar el resumen de la caracterización según tendencias.

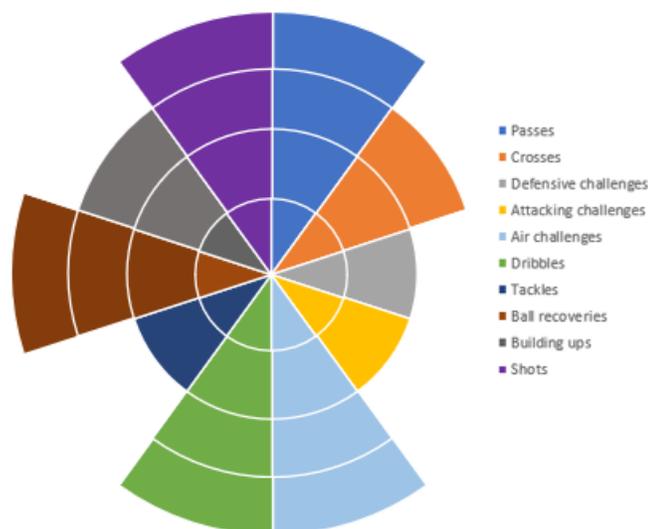


Figura 5.3: Comportamiento variables estilo 2

5.2.1.3. Estilo 3: Estilo de juego aéreo y de construcción

Se observa una tendencia promedio en tiros al arco. Se observa una baja tendencia en pases, lo que podría sugerir que en este estilo se tiende a evitar el juego de pases y, en su lugar, optar por enfoques más directos. Los centros tienen una tendencia promedio, no son la principal estrategia de ataque.

Baja tendencia en disputas defensivas ganadas y levemente alta tendencia en disputas ofensivas ganadas, lo que puede indicar una capacidad para mantener la posesión del balón en el área rival. El estilo muestra una gran fortaleza en las disputas aéreas, lo que sugiere que pueden aprovechar su altura y capacidad física para ganar duelos en el juego aéreo.

El equipo no suele utilizar regates ni barridas como una estrategia principal ya que hay baja tendencia en ambos. Recuperaciones de balón muestran una tendencia promedio. Alta tendencia en building ups lo que sugiere un enfoque en la construcción del juego.

En resumen, este estilo de juego se caracteriza por un enfoque en el juego aéreo, y una baja tendencia hacia los pases, regates y barridas. En la imagen 5.4 se puede observar el resumen de la caracterización según tendencias.

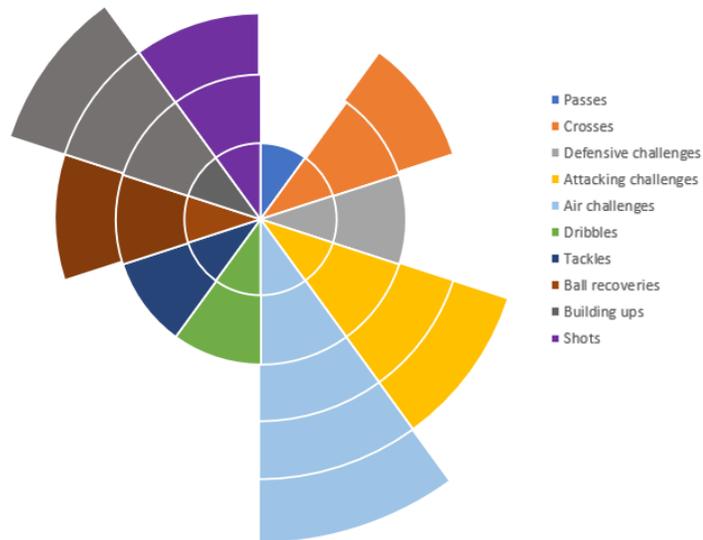


Figura 5.4: Comportamiento variables estilo 3

5.2.2. Cuatro estilos de juego

Se procede a describir el comportamiento de las variables para cada uno de los 4 estilos detectados. En la sección B del Anexo se pueden ver las distribuciones considerando más iteraciones en detalle. En la imagen 5.5 se puede observar las gráficas para cada estilo en comparación.

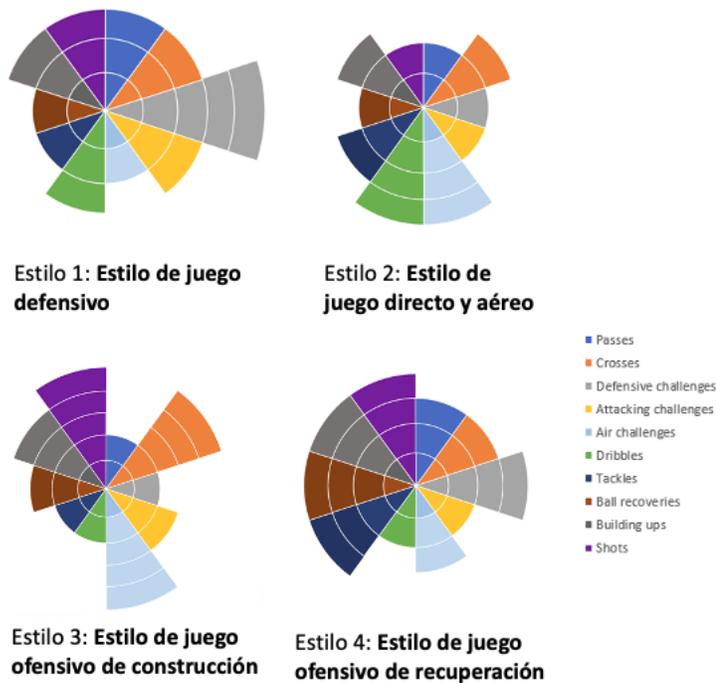


Figura 5.5: Comportamiento cuatro estilos

5.2.2.1. Estilo 1:Estilo defensivo

Este estilo de juego se caracteriza por un equilibrio en varios aspectos del juego. El estilo muestra una tendencia promedio en tiros, pases y centros, lo que indica que no se inclina demasiado hacia ninguna de estas acciones específicas. Sin embargo, donde realmente destacan es en las disputas defensivas ganadas, con una muy alta tendencia.

En cuanto a las disputas ofensivas ganadas, el estilo muestra una tendencia promedio, lo que indica una capacidad razonable para mantener la posesión del balón en situaciones ofensivas. Sin embargo, en las disputas aéreas ganadas, el equipo tiene una baja tendencia, lo que puede sugerir una menor fortaleza en el juego aéreo.

En términos de regates y barridas, muestran una tendencia promedio y baja, respectivamente, lo que indica que no dependen excesivamente de estas acciones. Además, el estilo presenta una levemente baja tendencia en recuperaciones de balón, lo que indica una menor eficiencia en ese ámbito. Por último, en términos de building ups, muestran una tendencia promedio.

En resumen, este estilo de juego se caracteriza por una defensa sólida, una capacidad razonable en disputas ofensivas, pero una baja tendencia en disputas aéreas ganadas y recuperaciones de balón. En la imagen 5.6 se puede observar el resumen de la caracterización según tendencias.



Figura 5.6: Comportamiento variables estilo 1

5.2.2.2. Estilo 2: Estilo directo y aéreo

El estilo de juego de este equipo en el fútbol se caracteriza por una tendencia baja en tiros al arco y pases, lo que indica que no buscan activamente crear oportunidades de gol a través de estos medios.

Sin embargo, mantienen una tendencia promedio en cuanto a los centros, lo que sugiere que aún consideran esta estrategia para atacar. En las disputas defensivas y ofensivas, muestran una baja tendencia de ganarlas, lo que podría indicar dificultades para recuperar el balón o superar a los oponentes en duelos individuales. Sin embargo, se destacan en las disputas aéreas, mostrando una alta tendencia de ganarlas, aprovechando su habilidad en el juego aéreo.

El estilo también muestra una inclinación hacia los regates, lo que implica que confían en la habilidad individual de sus jugadores para superar a los oponentes. En cuanto a las barridas y recuperaciones de balón, mantienen una tendencia promedio.

En general, su estilo se caracteriza por una tendencia promedio en los centros y las barridas, y destacando en el juego aéreo y los regates. En la imagen 5.7 se puede observar el resumen de la caracterización según tendencias.

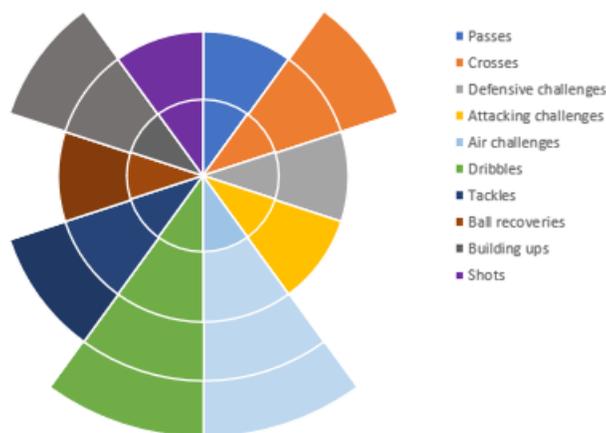


Figura 5.7: Comportamiento variables estilo 2

5.2.2.3. Estilo 3: Estilo ofensivo de construcción

Este estilo de juego se caracteriza por un enfoque ofensivo directo. El equipo muestra una muy alta tendencia en tiros al arco y centros, lo que indica una clara intención de buscar oportunidades de gol y atacar constantemente.

Muestran una levemente baja tendencia en pases. Además, tienen una alta tendencia en disputas aéreas ganadas, lo que demuestra su habilidad para aprovechar su altura y fortaleza física en el juego aéreo. A pesar de tener una baja tendencia en disputas defensivas ganadas, muestran una tendencia promedio en disputas ofensivas ganadas, lo que indica una capacidad razonable para mantener la posesión y generar oportunidades en el área contraria.

Los regates y barridas tienen una baja tendencia, lo que sugiere que el equipo no se basa en el juego individual o en acciones defensivas arriesgadas. En términos de recuperaciones de balón, muestran una tendencia promedio. Además, el equipo destaca por su alta tendencia en building ups, lo que indica una voluntad de construir juego.

En resumen, este estilo de juego se caracteriza por su enfoque ofensivo, con una alta tendencia en tiros, centros y disputas aéreas ganadas, y una preferencia por la construcción paciente del juego. En la imagen 5.8 se puede observar el resumen de la caracterización según tendencias.

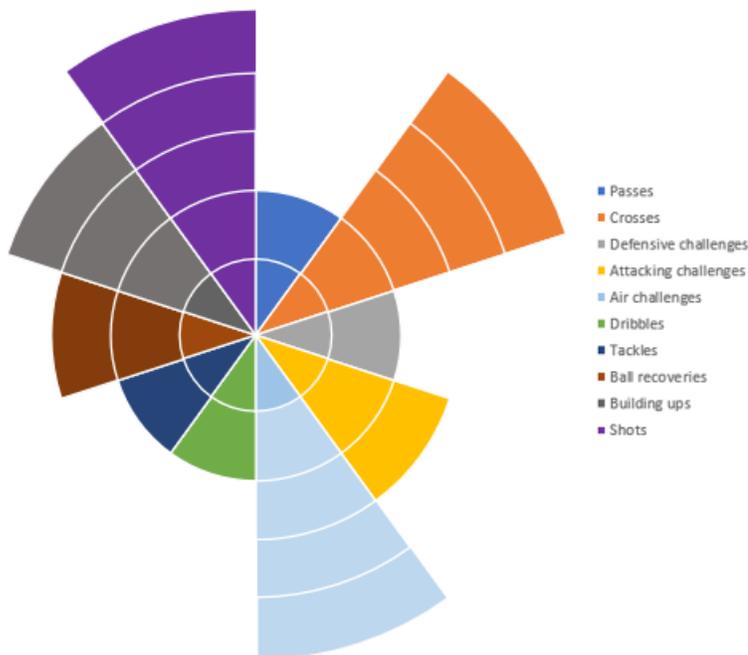


Figura 5.8: Comportamiento variables estilo 3

5.2.2.4. Estilo 4: Estilo ofensivo de recuperación

Este estilo de juego también se caracteriza por un enfoque ofensivo, con una alta tendencia en tiros al arco y una búsqueda constante de oportunidades de gol. Aunque los pases y centros se mantienen en una tendencia promedio, el equipo muestra una gran capacidad en la recuperación de balón, ganando disputas defensivas y realizando barridas con alta frecuencia.

A pesar de una tendencia levemente baja en disputas ofensivas ganadas, el equipo se destaca en la construcción del juego desde la defensa, con building ups que se encuentran levemente por encima del promedio. Para finalizar, muestra una baja tendencia en regates.

En general, este estilo de juego se caracteriza por un enfoque ofensivo, una sólida defensa y una mentalidad de recuperación de balón. En la imagen 5.9 se puede observar el resumen de la caracterización según tendencias.

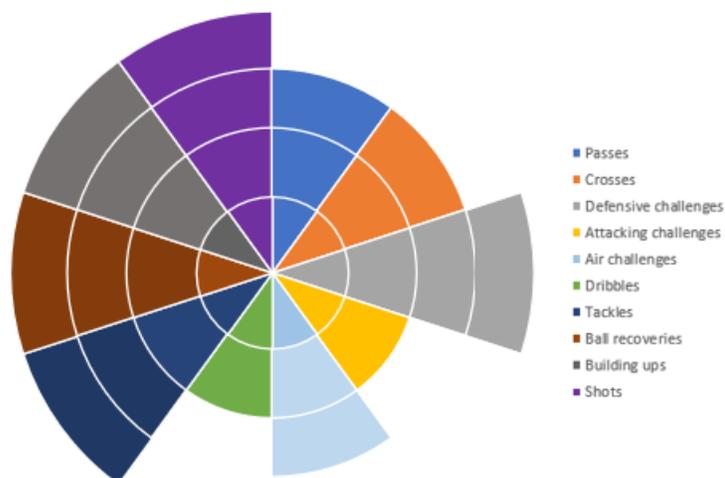


Figura 5.9: Comportamiento variables estilo 4

5.2.3. Cinco estilos de juego

Se procede a describir el comportamiento de las variables para cada uno de los 5 estilos detectados. En la sección C del Anexo se pueden ver las distribuciones con más iteraciones en detalle. En la imagen 5.10 se puede observar las gráficas para cada estilo en comparación.

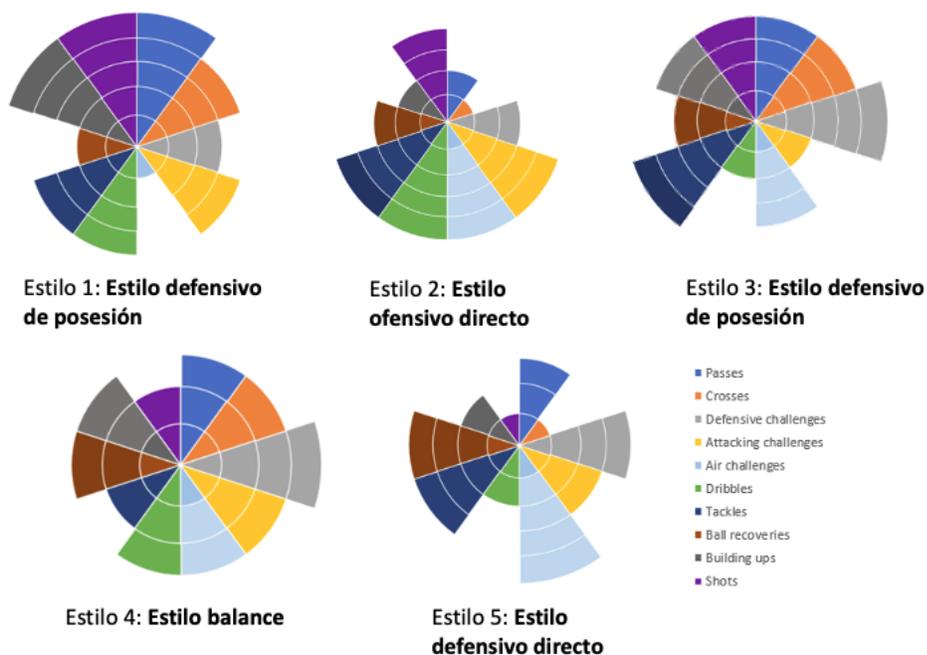


Figura 5.10: Comportamiento cinco estilos

5.2.3.1. Estilo 1: Estilo ofensivo de posesión

El estilo de juego se muestra ofensivo y orientado hacia el ataque. Tienen una alta tendencia de tiros al arco, lo cual indica que están constantemente buscando oportunidades de gol. Además, tiene alta tendencia de pases que sugiere mantener la posesión del balón. El estilo también muestra una alta tendencia de centros. Sin embargo, tienen una tendencia promedio en disputas defensivas ganadas y promedio en disputas aéreas ganadas.

En cuanto a las habilidades individuales, tienen una alta tendencia en regates y una alta tendencia en barridas, lo que sugiere que sus jugadores son agresivos en la recuperación del balón. Por otro lado, tienen una baja tendencia en recuperaciones de balón, lo que indica que pueden tener dificultades en la fase de defensa y en la presión sobre el equipo contrario. Sin embargo, muestran una alta tendencia en building ups.

En resumen, el estilo de juego se caracteriza por un enfoque ofensivo, con énfasis en los tiros al arco, los pases y los centros. En la imagen 5.11 se puede observar el resumen de la caracterización según tendencias.

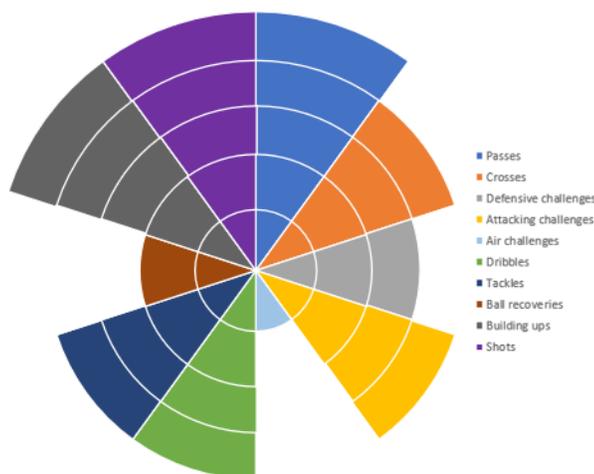


Figura 5.11: Comportamiento variables estilo 1

5.2.3.2. Estilo 2: Estilo ofensivo directo

El estilo muestra una baja tendencia en pases, lo que implica que no depende mucho del juego de pases para avanzar en el campo. El estilo también, muestra una alta tendencia en regates y barridas, lo que podría indicar que se sienten cómodos y confiados en el juego individual y están dispuestos a asumir riesgos para recuperar el balón o evitar que el rival avance.

Tiene una alta tendencia de tiros al arco, lo que indica que el equipo busca constantemente oportunidades de ataque y muestra una mentalidad ofensiva. Tiene una baja tendencia en centros, lo que sugiere que no depende mucho de estos.

En términos defensivos, tiene una baja tendencia en disputas defensivas ganadas, lo que implica que puede tener dificultades para recuperar el balón y defender efectivamente. Por otro lado, muestra una alta tendencia en disputas ofensivas ganadas y disputas aéreas ga-

nadas, lo que sugiere que son fuertes en duelos individuales y en el juego aéreo. Por último, tienen una baja tendencia en recuperaciones de balón y building ups.

En resumen, muestra un estilo de juego ofensivo y directo, con una preferencia por acciones individuales, tiros al arco y juego físico. Sin embargo, pueden tener dificultades en la fase defensiva y en la elaboración del juego. En la imagen 5.12 se puede observar el resumen de la caracterización según tendencias.

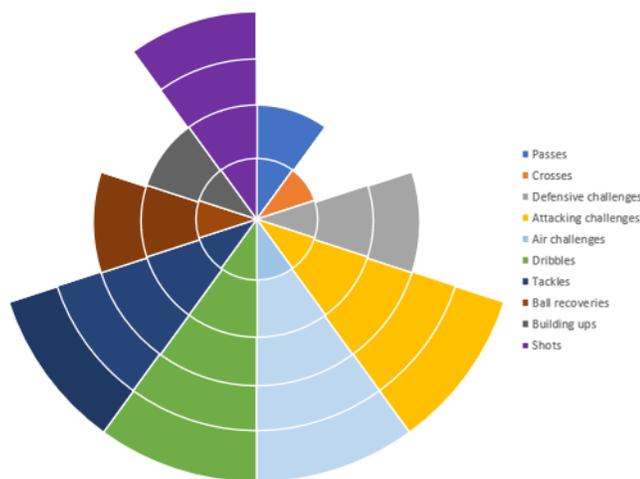


Figura 5.12: Comportamiento variables estilo 2

5.2.3.3. Estilo 3: Estilo defensivo de posesión

En disputas defensivas, muestran una alta tendencia en ganarlas, por encima del promedio. Esto indicaría que el equipo tiene una sólida defensa y es capaz de recuperar el balón en situaciones defensivas. Sin embargo, en disputas ofensivas, muestran una baja tendencia. Esto, podría indicar que el equipo tiene dificultades para mantener la posesión del balón en situaciones ofensivas o puede tener problemas para superar la defensa contraria en duelos individuales. También, muestra una alta tendencia en barridas, muy por encima del promedio. En términos de recuperaciones de balón, muestran una tendencia promedio. Por último, muestra una alta tendencia en building ups.

El estilo tiene una alta tendencia de pases. También, muestran una alta tendencia de tiros al arco, lo que indica que son ofensivos y buscan constantemente oportunidades para disparar al arco rival. En cuanto a los centros, el equipo muestra una tendencia promedio, lo que indica que no se basan excesivamente en los centros como parte de su estrategia de juego. El estilo también muestra una alta tendencia en ganar disputas aéreas. En cuanto a los regates, el equipo muestra una baja tendencia.

En resumen, el estilo de juego del equipo se caracteriza por una alta tendencia de pases, tiros al arco, disputas defensivas ganadas, disputas aéreas ganadas, barridas y building ups. Sin embargo, presentan baja tendencia en disputas ofensivas ganadas y regates. En la imagen 5.13 se puede observar el resumen de la caracterización según tendencias.

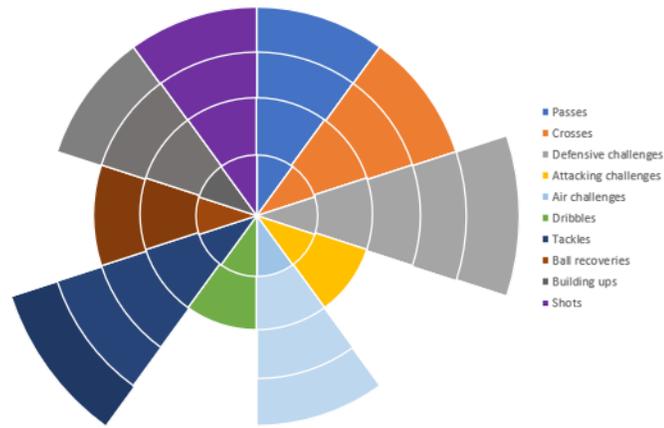


Figura 5.13: Comportamiento variables estilo 4

5.2.3.4. Estilo 4: Estilo equilibrado

El equipo muestra valores promedio en términos de pases, centros, disputas aéreas ganadas, disputas ofensivas ganadas, regates y building ups, ya que todas estas características se sitúan alrededor del promedio. Esto sugiere que el equipo no muestra una inclinación notable hacia ninguna de estas áreas en particular.

Sin embargo, hay algunas áreas en las que el equipo muestra una baja tendencia. Estas áreas incluyen tiros al arco y barridas. Esto puede indicar que el estilo puede ser menos agresivo en términos de recuperación de balones y defensa.

En resumen, el estilo de juego del equipo se caracteriza por un enfoque equilibrado en ámbitos defensivos y ofensivos. En la imagen 5.14 se puede observar el resumen de la caracterización según tendencias.

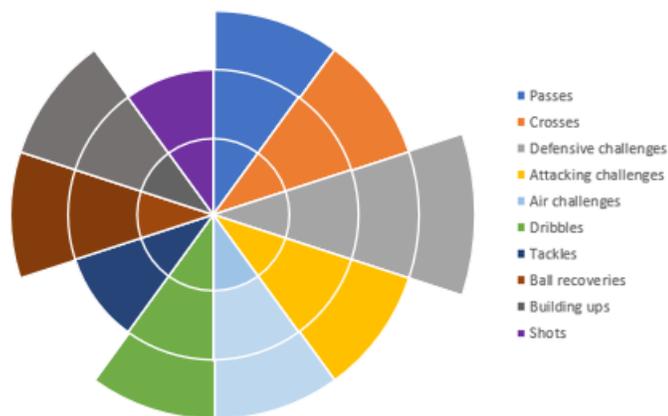


Figura 5.14: Comportamiento variables estilo 4

5.2.3.5. Estilo 5: Estilo defensivo directo

El estilo tiene una baja tendencia en tiros al arco y centros, lo que indica que no se enfocan en crear muchas oportunidades de gol a través de disparos o centros al área. También, tienen una baja tendencia en regates, lo que sugiere que no dependen mucho de la habilidad individual de los jugadores para superar a los oponentes. En cuanto a los pases, el equipo muestra una tendencia media.

El estilo tiene una alta tendencia en disputas defensivas y promedio en ofensivas ganadas. Sin embargo, el equipo muestra una alta tendencia en disputas aéreas ganadas, lo que indica que son fuertes en el juego aéreo.

También, tienen una alta tendencia en barridas, lo que sugiere que son agresivos en el juego defensivo y no dudan en realizar barridas para recuperar el balón. Por último, el equipo muestra una alta tendencia en recuperaciones de balón.

En general, el estilo de juego del equipo se caracteriza por ser más defensivo y táctico, con énfasis en disputas aéreas, barridas y recuperaciones de balón. En la imagen 5.14 se puede observar el resumen de la caracterización según tendencias.



Figura 5.15: Comportamiento variables estilo 5

En anexo están disponibles las imágenes de la distribuciones por cada una de las variables para la clasificación de 5 estilos.

Para efectos de los siguientes análisis, se escogió la clasificación de 5 estilos dado a que es la que cuenta con los estilos más diferenciados entre ellos y la que entrega más información.

5.3. Asignación de estilos de juego

Tras aplicar el modelo y analizar los datos recopilados, se han obtenido resultados reveladores sobre las probabilidades de cada equipo de elegir un estilo de juego determinado. Estos resultados ofrecen una visión enriquecedora de las estrategias empleadas, permitiendo identificar patrones y tendencias que se manifiestan a lo largo de la competición. A continuación, se

presentaran los hallazgos más destacados del análisis, brindando una comprensión profunda del estilo de juego adoptado por cada equipo y su influencia en los resultados obtenidos.

Se decide asignar 5 estilos de juego para las 10 variables anteriormente descritas. Dado los estilos de juego identificados en el modelo, es factible analizar a cada equipo y partido de la investigación. Utilizando los datos de parámetros de estilo de la última iteración del modelo es posible calcular el estilo de juego que usará un equipo dado su rendimiento en los partidos. Considerando 5 estilos de juego, cada uno representado por un vector de rendimiento. Cada vector tiene 10 valores numéricos que describen el rendimiento promedio de los estilos.

Hay una matriz de rendimiento donde cada fila representa el rendimiento de un equipo. Cada equipo tiene valores numéricos para las mismas 10 variables que se miden en los estilos de juego. Para calcular las probabilidades de que un equipo juegue con cada uno de los 5 estilos de juego, se utiliza la distancia euclidiana entre el rendimiento del equipo y los rendimientos de los estilos de juego. La distancia euclidiana mide qué tan similar o cercano es el rendimiento del equipo en comparación con cada estilo de juego.

Se toma la diferencia entre el rendimiento del equipo y el rendimiento de cada estilo de juego, calculando la norma (longitud) del vector resultante. Luego, se calcula la probabilidad de jugar con cada estilo de juego tomando el inverso de la distancia euclidiana. Una distancia menor corresponde a una probabilidad mayor. Las probabilidades se normalizan dividiendo cada probabilidad por la suma de todas las probabilidades.

El análisis realizado permite asignar los estilos con cierta probabilidad y determinar cuál es el estilo de juego predominante para cada equipo. Al asignar un estilo mayoritario a cada equipo, se obtiene una visión general de las preferencias y tendencias de juego de los equipos analizados, lo que puede ser de gran utilidad para comprender y analizar su desempeño en la liga.

En la tabla 5.1 se puede observar un ejemplo de como se observan los resultados. Para el primer partido de Curicó Unido se puede observar que existen dos estilos predominantes, esto podría indicar cambios de estilo a lo largo del partido. En cambio, para el partido de Melipilla se puede ver un estilo predominante.

Tabla 5.1: Probabilidad de que en un partido cada equipo juegue con cada uno de los estilos

ID	P(E1)	P(E2)	P(E3)	P(E4)	P(E5)
Equipo 1 (CUR) Partido 1	0.1	0.1	0.05	0.31	0.44
Equipo 2 (MEL) Partido 1	0.45	0.08	0.11	0.21	0.15

Al observar el comportamiento de un sólo equipo en el torneo, por ejemplo Universidad Católica en la tabla 5.2, se puede notar que existen estilos predominantes a lo largo del torneo, y que si bien existen partidos donde hay dos estilos que destacan por sobre los otros, hay partidos en lo que existe solo un estilo predominante. Análisis más profundos se pueden hacer considerando más variables como por ejemplo, el oponente, para ver si recurre a estilos más defensivos u ofensivos según cual es el equipo contrario. El mismo ejercicio puede ser repetido para cada uno de los equipos del torneo.

Tabla 5.2: Probabilidad de que Universidad Católica juegue con cada uno de los estilos en 14 partidos del torneo

ID equipo	ID partido	Probabilidades por cada estilo	Estilo designado
13	7	[0.19, 0.09, 0.08, 0.42, 0.23]	4
13	14	[0.19, 0.09, 0.08, 0.42, 0.22]	4
13	18	[0.43, 0.08, 0.12, 0.22, 0.15]	1
13	26	[0.1, 0.09, 0.05, 0.39, 0.35]	4
13	46	[0.53, 0.07, 0.13, 0.16, 0.12]	1
13	48	[0.22, 0.09, 0.09, 0.39, 0.21]	4
13	61	[0.53, 0.07, 0.12, 0.16, 0.12]	1
13	68	[0.28, 0.08, 0.37, 0.14, 0.12]	3
13	73	[0.54, 0.07, 0.12, 0.16, 0.12]	1
13	83	[0.09, 0.14, 0.05, 0.26, 0.45]	5
13	95	[0.37, 0.09, 0.11, 0.25, 0.17]	1
13	100	[0.26, 0.09, 0.1, 0.35, 0.2]	4
13	107	[0.14, 0.08, 0.06, 0.5, 0.22]	4
13	115	[0.24, 0.09, 0.09, 0.36, 0.21]	4
13	128	[0.1, 0.17, 0.06, 0.24, 0.43]	5

Otro análisis interesante que se puede hacer es ver que tipos de estilos compiten entre sí y con que frecuencia. Asumiendo que el estilo de juego que se enfrentó, es el con mayor probabilidad podemos observar cuales son los que más compiten entre ellos. Se puede observar en la imagen 5.16 que los estilos que más competirían entre sí son el 4 y el 5.

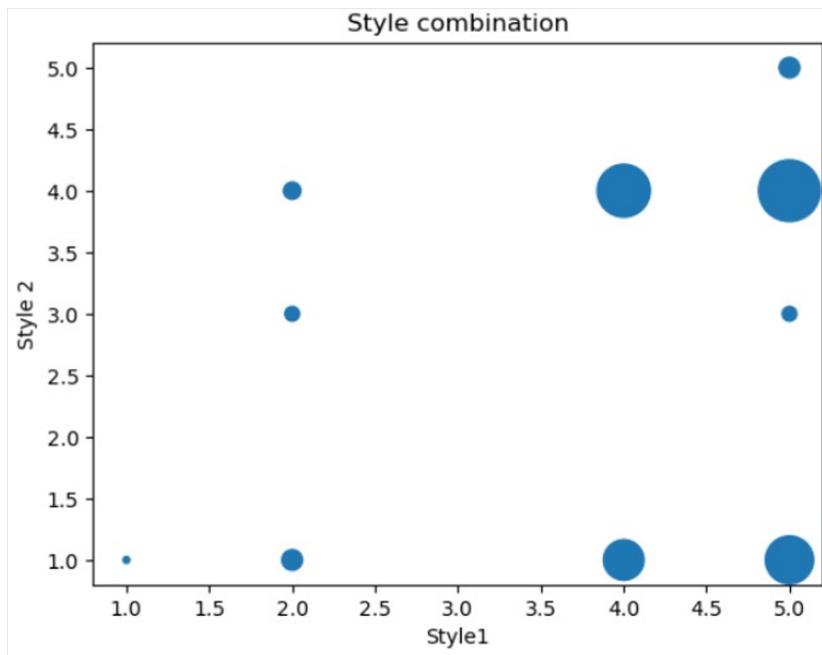


Figura 5.16: Combinaciones entre estilos más repetidas

En resumen, este análisis representa solo una parte de las posibles investigaciones y análisis que se pueden realizar sobre los datos disponibles. Hay una amplia gama de enfoques y técnicas que se pueden aplicar para profundizar en el estudio de los resultados.

El objetivo de esta investigación se ha cumplido al detectar diferentes estilos de juego dentro del Torneo Nacional Chileno. Sin embargo, es importante destacar que este es solo el comienzo y que hay muchas oportunidades para seguir investigando y obteniendo nuevos conocimientos a partir de estos datos.

Capítulo 6

Conclusiones

Esta investigación utiliza datos de la primera división del fútbol chileno año 2021 a través del proveedor Instat. Para cada una de las fechas se consideran variables que describen de forma agregada lo que sucedió en la cancha a nivel partido y por equipos tales como: disputas defensivas, disputas ofensivas, pases, tiros al arco, regates, etc. En el análisis exploratorio de datos, se pudo determinar qué variables eran más adecuadas para el modelo.

El modelo creado, utiliza inferencia bayesiana para asignar estilos de juego a los equipos de fútbol. Utiliza datos de observaciones 10 para estimar los parámetros de 5 estilos de juego y las probabilidades de estilos de juego mediante distribuciones de prior y verosimilitud.

La presente investigación ha logrado cumplir con su objetivo principal, que correspondía a detectar estilos de juego presentes en los equipos de fútbol profesional chileno. A través de un exhaustivo análisis de datos y un enfoque metodológico sólido, se ha examinado y categorizado de manera precisa los diferentes patrones de juego utilizados por los equipos en la liga correspondiente al año 2021. En particular se decidió analizar en profundidad la clasificación de 5 estilos.

Esta investigación aborda un tema de gran relevancia al utilizar la perspectiva de la inferencia bayesiana, un enfoque metodológico innovador en clasificación de estilos de juego. Es notable la ausencia de investigaciones similares en el contexto del fútbol latinoamericano, especialmente en el ámbito chileno. Esta combinación única de enfoques metodológicos y geográficos contribuye al conocimiento del fútbol chileno y proporciona información valiosa para el desarrollo de futuras investigaciones, así como para la toma de decisiones estratégicas en el ámbito deportivo y la mejora del rendimiento de los equipos.

Existen varias oportunidades de mejora para el modelo creado, a considerar para futuras iteraciones. En primer lugar, sería beneficioso mejorar las variables utilizadas, considerando otras que describan quizás de manera más precisa el desarrollo del partido. Algunas de las limitaciones de este modelo están dadas por los datos con los que se cuenta.

Una opción interesante a explorar es considerar un partido como sólo una observación, en la que cada equipo en cada partido usa un estilo con cierta probabilidad considerando a quien se enfrenta y de esta forma generar estrategias. Mientras se calibra el modelo se buscaría llegar a un especie de equilibrio.

Por último, ampliar el alcance del modelo incluyendo más equipos o ligas más grandes, permitiría obtener una visión más completa y generalizable. Por último, contar con una mayor cantidad de datos, es decir, más partidos, sería fundamental para ajustar y perfeccionar el modelo, ya que una mayor variedad de situaciones y escenarios aumentaría su capacidad de clasificar.

Bibliografía

- Adam Hewitt, G. G. . K. N. (2016, 03). Game style in soccer: what is it and can we quantify it? *International Journal of Performance Analysis in Sport*, 16:1, 355-372. doi: 10.1080/24748668.2016.11868892
- Baio, G., y Blangiardo, M. (2010, 02). Bayesian hierarchical model for the prediction of football results. *Journal of Applied Statistics*, 37, 253-264. doi: 10.1080/02664760802684177
- Deshpande, S., y Wyner, A. (2017). A hierarchical bayesian model of pitch framing. *Journal of Quantitative Analysis in Sports*, 13(3), 95–112.
- Drezner, R., Lamas, L., Barrera, J., y Dantas, L. (2020, 05). Original article a method for classifying and evaluating the efficiency of offensive playing styles in soccer. *Journal of Physical Education and Sport*, 20, 1284-1294. doi: 10.7752/jpes.2020.03179
- Fernández Navarro, F. J. (2019). Analysis of styles of play in soccer and their effectiveness. *Granada: Universidad de Granada*. doi: <http://hdl.handle.net/10481/54554>
- Gregory, S. (2019). Ready player run: Off-ball run identification and classification. *2019 Barcelona Sports Analytics Conference*. doi: https://static.capabiliaserver.com/frontend/clients/barca/wp_prod/wp-content/uploads/2020/01/ed15d067-ready-player-run-barcelona-paper-sam-gregory.pdf
- Jeroen Clijmans, M. V. R. y. J. D. (2023). Looking beyond the past: Analyzing the intrinsic playing style of soccer teams. *Lecture Notes in Computer Science book series*, 13718. doi: <https://punditarena.com/football/paul-moore/fernando-torres-chelsea/>
- Jurek, A., y Zakrzewska, D. (2008). Improving naïve bayes models of insurance risk by unsupervised classification. , 137-144. doi: 10.1109/IMCSIT.2008.4747230
- Karlis, D., y Ntzoufras, I. (2008, 08). Bayesian modelling of football outcomes: Using the skellam’s distribution for the goal difference. *IMA Journal of Management Mathematics - IMA J MANAG MATH*, 20. doi: 10.1093/imaman/dpn026
- Maddox, J., Sides, R., y Harvill, J. (2022). Bayesian estimation of in-game home team win probability for college basketball. *Journal of Quantitative Analysis in Sports*, 18(3), 201–213. doi: 10.1515/jqas-2021-0086
- Moore, P. (2020). Fernando torres explains why his form fell to pieces at chelsea. *Pundit Arena*. doi: <https://punditarena.com/football/paul-moore/fernando-torres-chelsea/>
- Plakias S, K. C. T. T. P. M. P. D. G. G. T. D., Moustakidis S. (2023, 03). Identifying soccer teams styles of play: A scoping and critical review. *Journal of Functional Morphology and Kinesiology.*, 8(2):39. doi: <https://doi.org/10.3390/jfmk8020039>
- Tom L. G. Bergkamp, A. S. M. N. R. R. M. . R. J. R. d. H., Wouter G. P. Frencken. (2022). How soccer scouts identify talented players. *European Journal of Sport Science*, 22:7, 994-1004. doi: <https://www.tandfonline.com/doi/full/10.1080/17461391.2021.1916081>
- Whitaker, G. A., Silva, R., Edwards, D., y Kosmidis, I. (2017). A bayesian approach for determining player abilities in football. *arXiv, Cornell University*. doi: <https://arxiv.org/pdf/1710.00001.pdf>

Anexos

Anexo A. Distribuciones de cada variable con 3 estilos

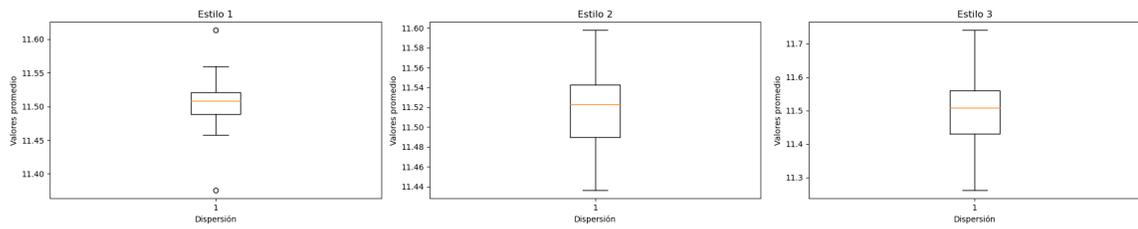


Figura A.1: Gráfico comparación parámetro Shots

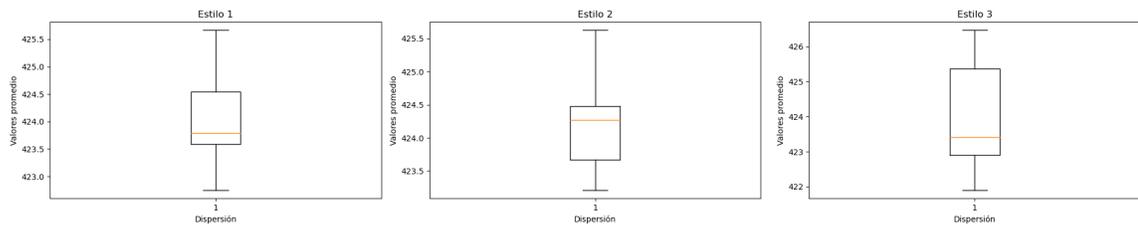


Figura A.2: Gráfico comparación parámetro Passes

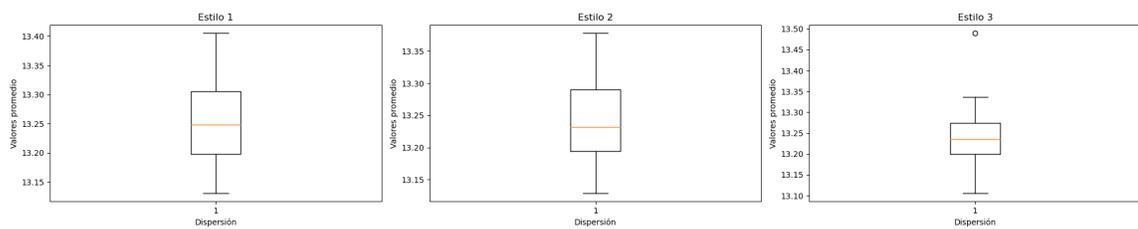


Figura A.3: Gráfico comparación parámetro Crosses

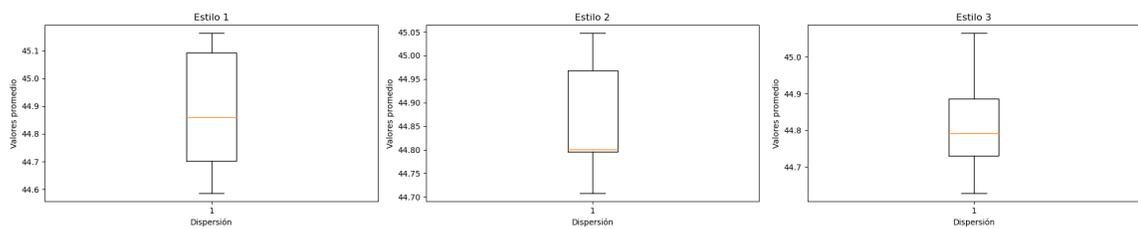


Figura A.4: Gráfico comparación parámetro Defensive challenges won

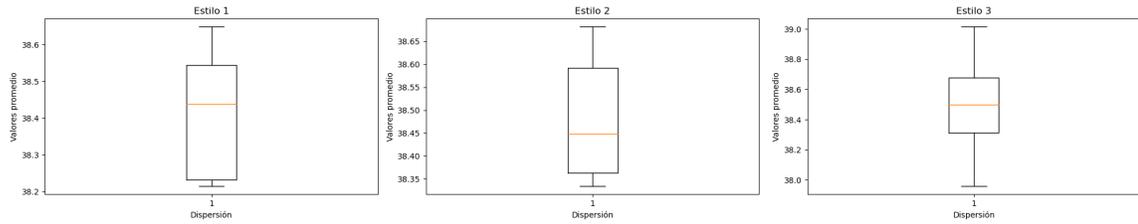


Figura A.5: Gráfico comparación parámetro Attacking challenges won

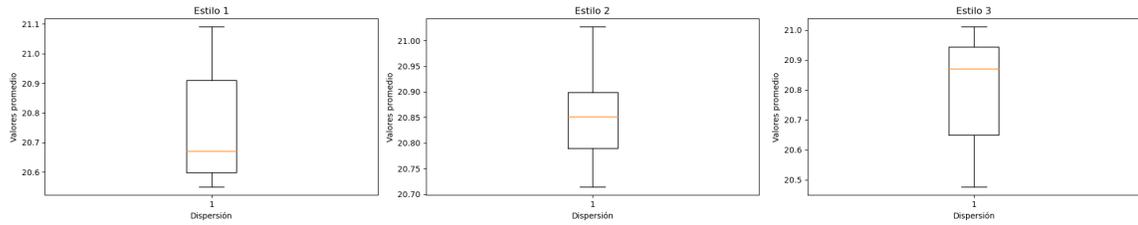


Figura A.6: Gráfico comparación parámetro Air challenges won

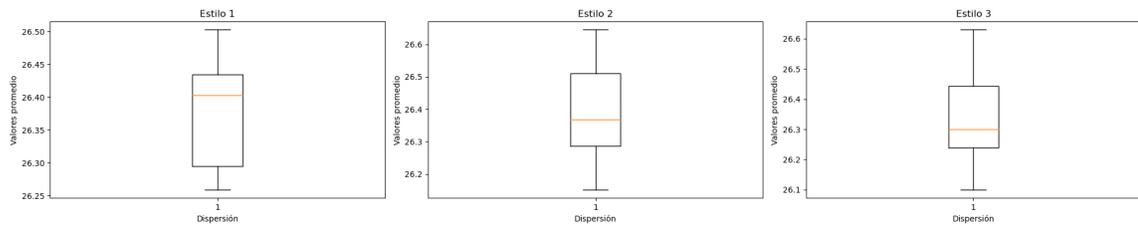


Figura A.7: Gráfico comparación parámetro Dribbles

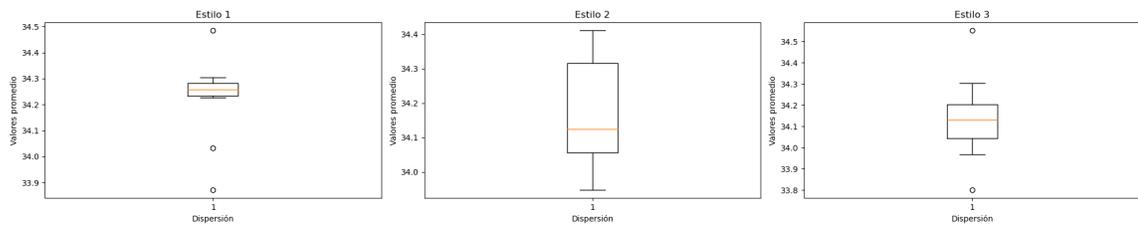


Figura A.8: Gráfico comparación parámetro Tackles

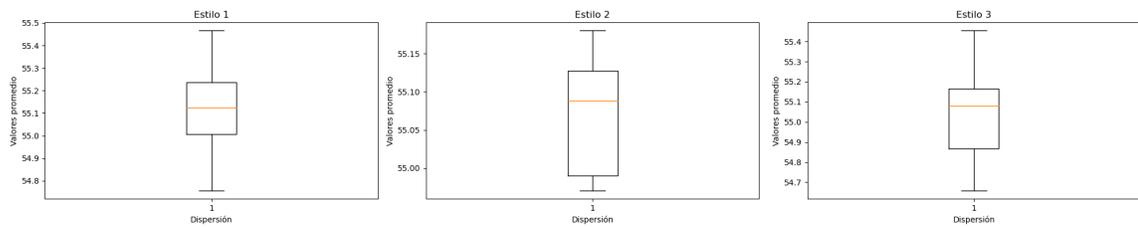


Figura A.9: Gráfico comparación parámetro Ball Recoveries

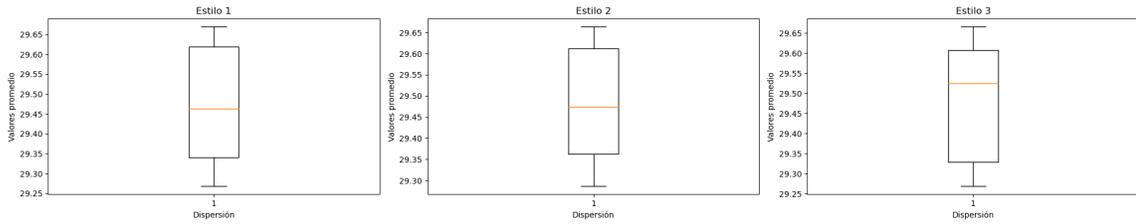


Figura A.10: Gráfico comparación parámetro Building ups

Anexo B. Distribuciones de cada variable con 4 estilos

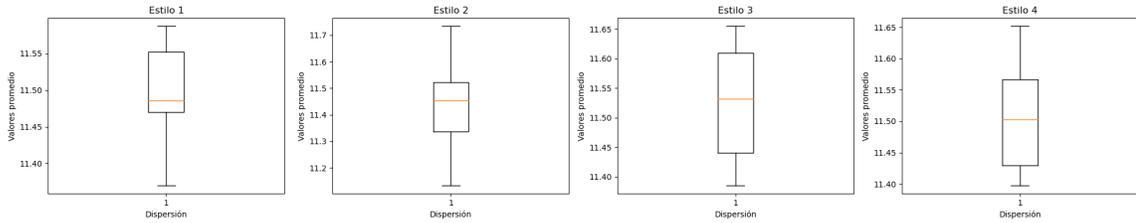


Figura B.1: Gráfico comparación parámetro Shots

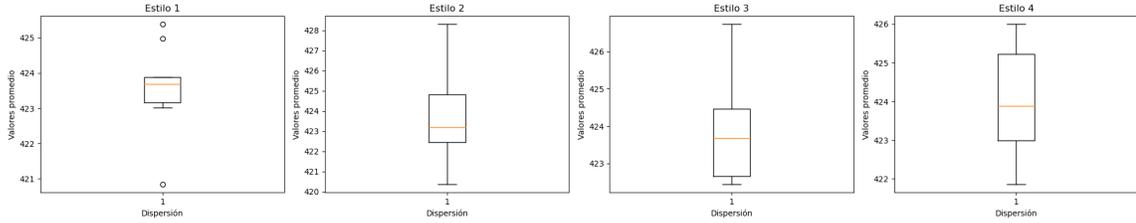


Figura B.2: Gráfico comparación parámetro Passes

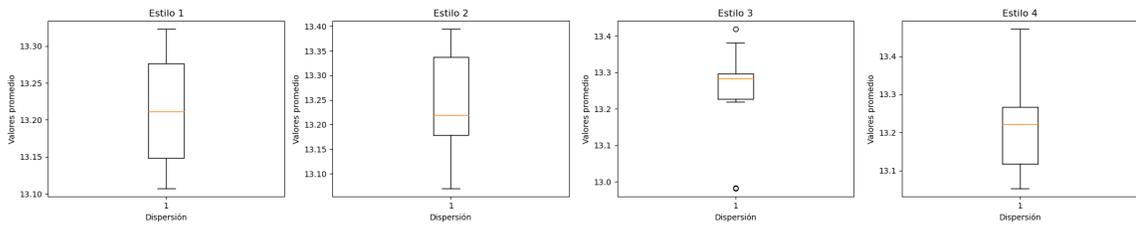


Figura B.3: Gráfico comparación parámetro Crosses

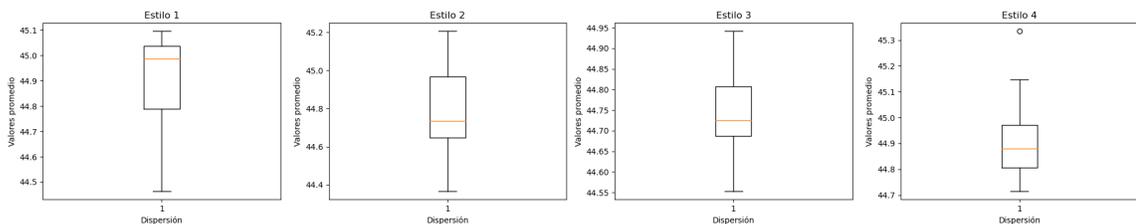


Figura B.4: Gráfico comparación parámetro Defensive challenges won

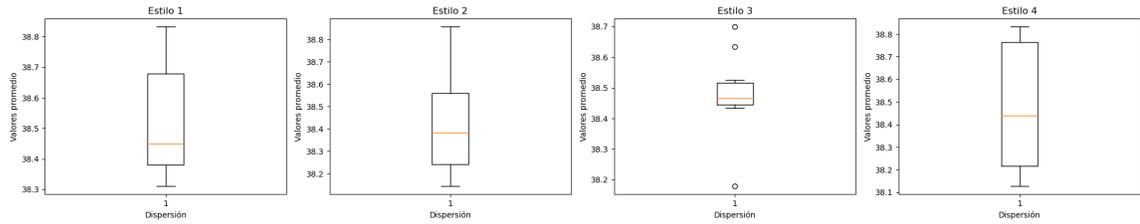


Figura B.5: Gráfico comparación parámetro Attacking challenges won

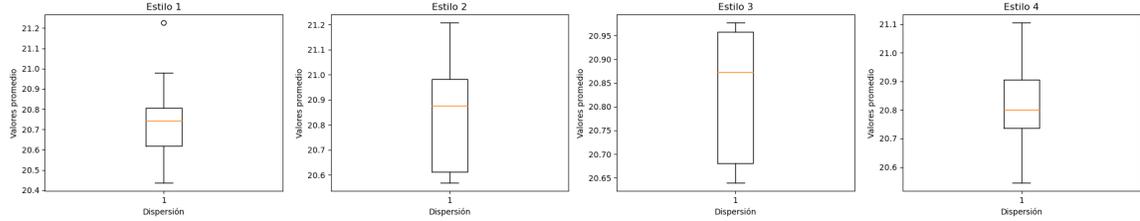


Figura B.6: Gráfico comparación parámetro Air challenges won

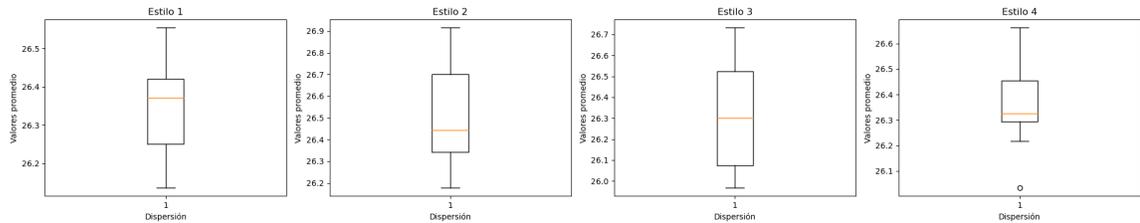


Figura B.7: Gráfico comparación parámetro Dribbles

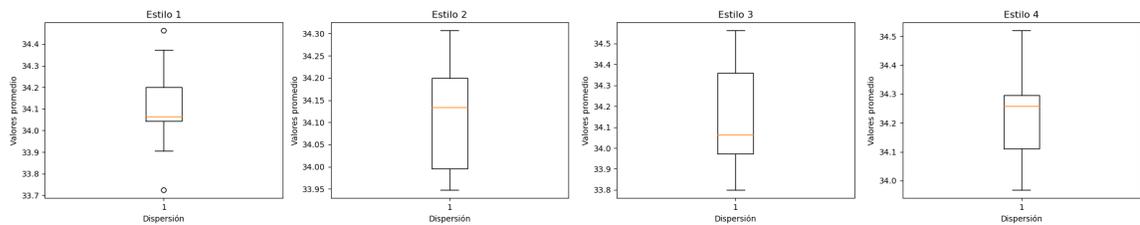


Figura B.8: Gráfico comparación parámetro Tackles

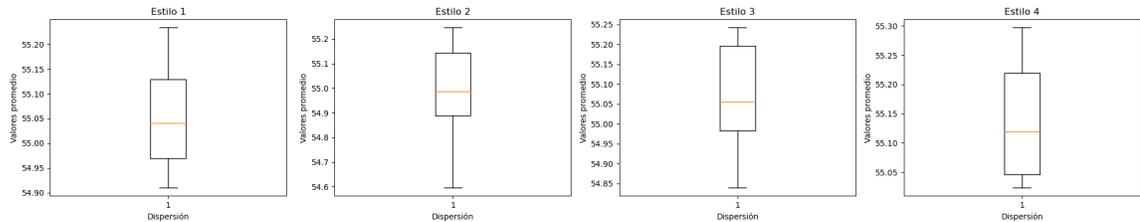


Figura B.9: Gráfico comparación parámetro Ball Recoveries

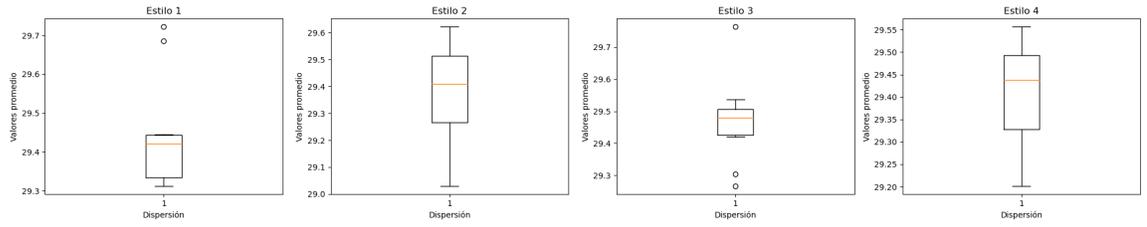


Figura B.10: Gráfico comparación parámetro Building ups

Anexo C. Distribuciones de cada variable con 5 estilos

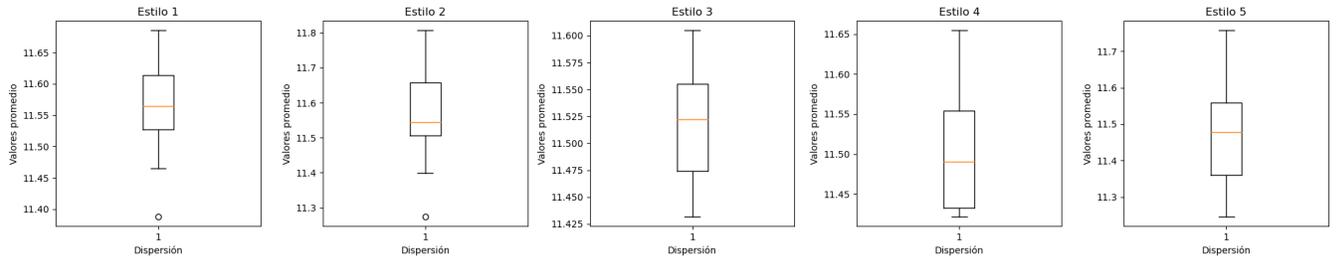


Figura C.1: Gráfico comparación parámetro Shots

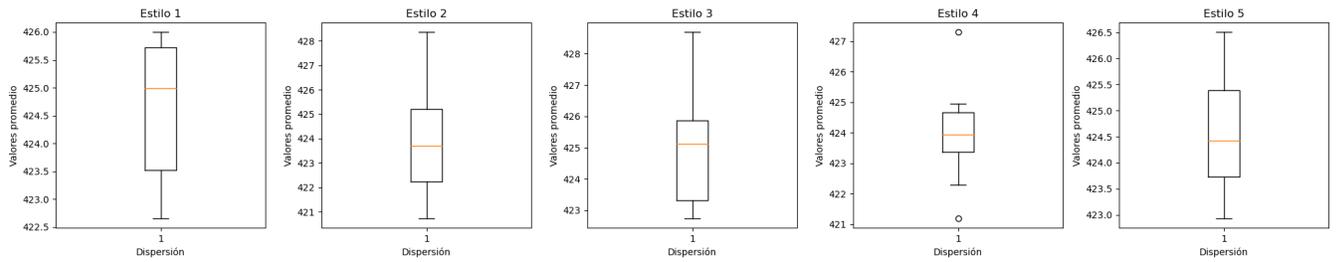


Figura C.2: Gráfico comparación parámetro Passes

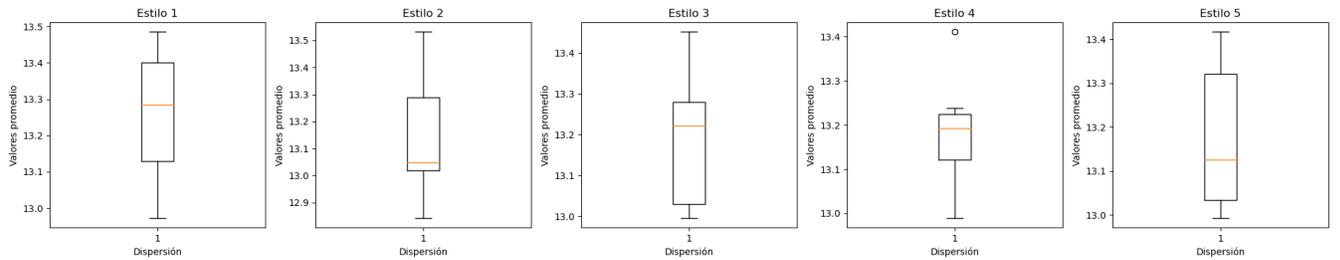


Figura C.3: Gráfico comparación parámetro Crosses

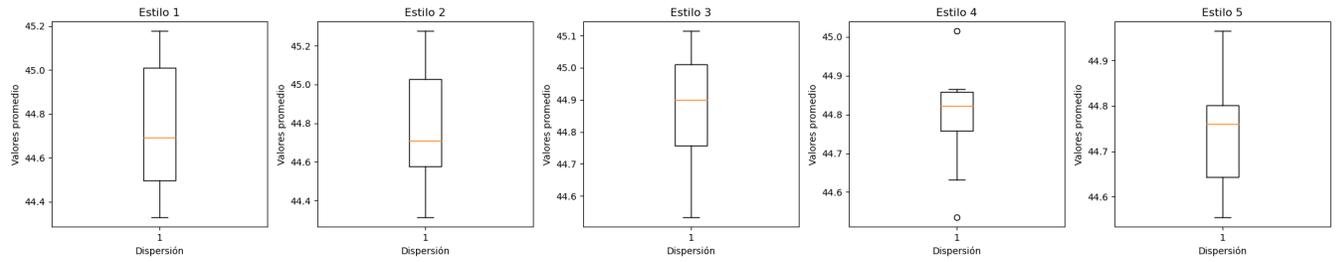


Figura C.4: Gráfico comparación parámetro Defensive challenges won

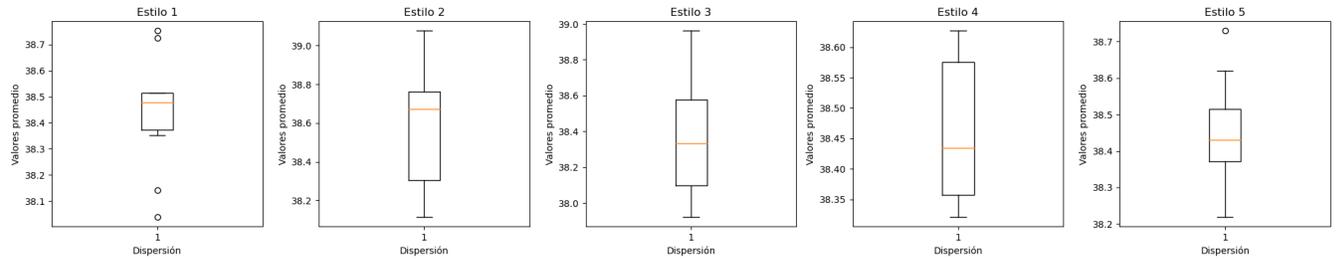


Figura C.5: Gráfico comparación parámetro Attaking challenges won

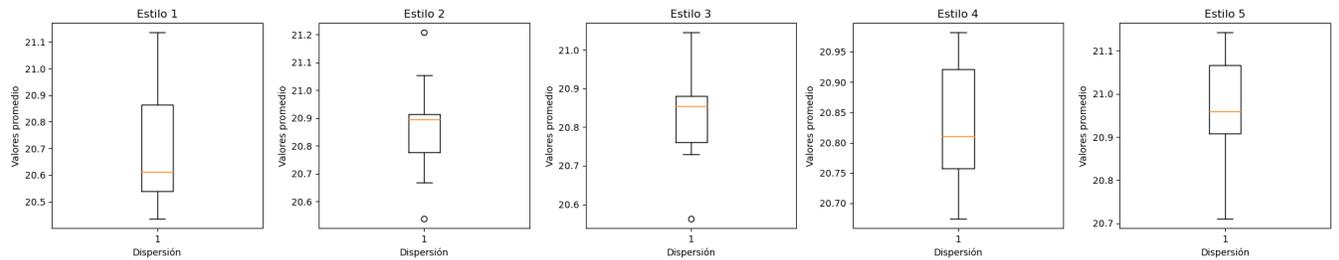


Figura C.6: Gráfico comparación parámetro Air challenges won

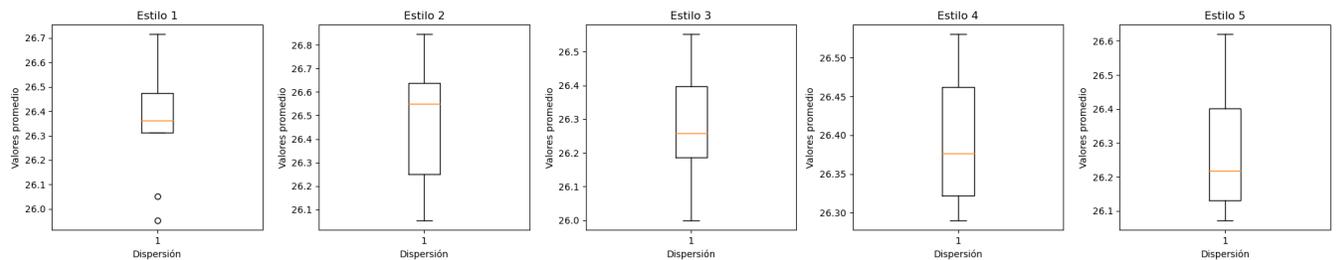


Figura C.7: Gráfico comparación parámetro Dribbles

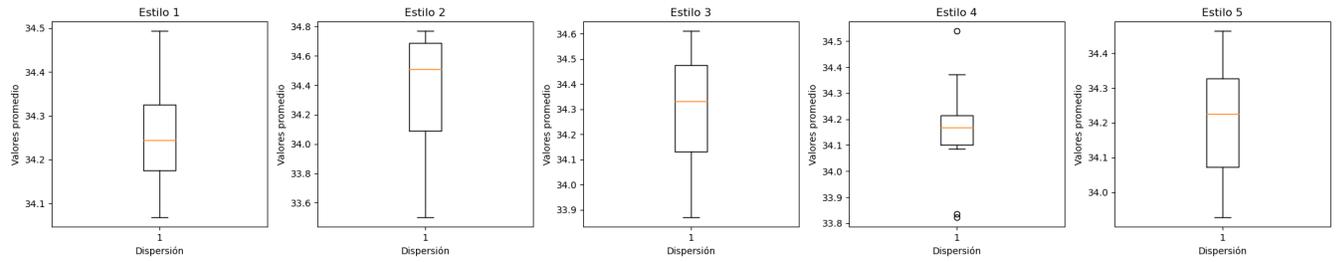


Figura C.8: Gráfico comparación parámetro Tackles

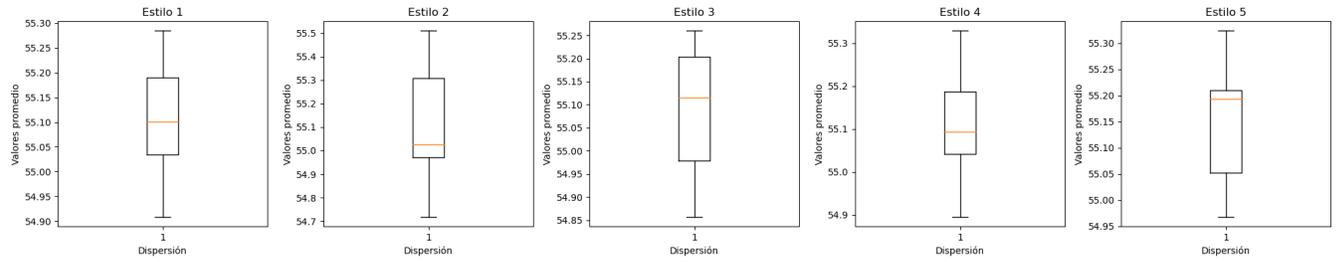


Figura C.9: Gráfico comparación parámetro Ball Recoveries

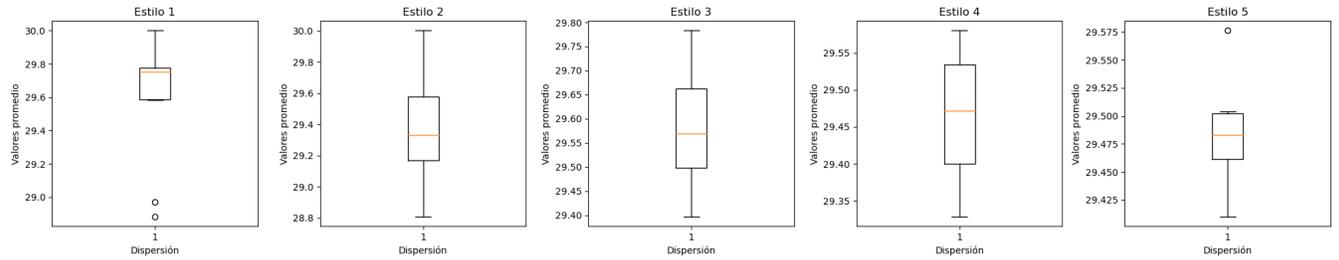


Figura C.10: Gráfico comparación parámetro Building ups